2019

# Genetics and Molecular Dynamics of L,L-Diaminopimelate Aminotransferase (DapL): an Enzyme Involved in Lysine and Peptidoglycan Biosynthesis

Lily E.J. Adams
lea4012@rit.edu

# Genetics and Molecular Dynamics of L,L-Diaminopimelate Aminotransferase (DapL): an Enzyme Involved in Lysine and Peptidoglycan Biosynthesis

By: Lily E.J. Adams

Submitted in Partial Fulfillment of the Requirements for the Master of Science Degree in Bioinformatics at The Rochester Institute of Technology

2019

**Rochester Institute of Technology**
**Thomas H. Gosnell School of Life Sciences**
**Bioinformatics Program**

**To:** Head, Thomas H. Gosnell School of Life Sciences

The undersigned state that Lily Elizabeth Julia Adams, a candidate for the Master of Science degree in Bioinformatics, has submitted her thesis and has satisfactorily defended it.

This completes the requirements for the Master of Science degree in Bioinformatics at Rochester Institute of Technology.

**Thesis committee members:**

| **Name** | **Date** |
|---|---|
| André O. Hudson, Ph.D.<br>Thesis Advisor | |
| Gregory A. Babbitt, Ph.D. | |
| Michael A. Savka, Ph.D | |

# 1 ABSTRACT

The marked increase of bacteria that are resistant to clinically relevant antibiotics in recent years has sparked a push in the development and discovery of novel antibiotics. Complementary to this, the identification of novel antibiotic targets is also of great interest. The lysine biosynthesis pathway is an ideal candidate for these efforts because its product is an essential amino acid. Lysine cannot be synthesized by humans and thus poses minimal risk of side effects if synthesis is targeted. The penultimate product in the lysine biosynthesis pathway, *meso*-DAP, is also a key component of the gram-negative bacterial cell wall peptidoglycan layer. The L,L-diaminopimelate aminotransferase (DapL) pathway, a recently discovered variant in the lysine biosynthetic pathway is especially of interest because it has been identified in only about 13% of bacteria. Because of its narrow distribution in bacteria, it could be an ideal target for the development of narrow spectrum antibiotics. A key enzyme in the pathway, DapL is a homodimer that catalyzes the conversion of tetrahydrodipicolinate (THDP) to L,L-diaminopimelate (L,L-DAP) in a single, reversible transamination reaction. While this enzyme is essential in plants that contain the pathway, it is not directly known whether the same is true for bacteria that contain the pathway. In order to evaluate DapL as a target for the development of narrow spectrum antibiotics, genetic and phenotypic characterizations must be performed. This project evaluated DapL from *Verrucomicrobium spinosum*, a close relative of Chlamydia, as a potential target in the development of antibiotic compounds. It assessed the essentiality of the *VsdapL* gene via targeted knockouts using the CRISPR/cas9 system, providing evidence supporting the essentiality of the gene in bacteria. It also evaluated putative antagonistic ligands using a comprehensive and comparative molecular dynamics (MD) software package – DROIDS (Detecting Relative Outlier Impacts in Dynamic Simulations) 2.0, providing insight into the dynamic behavior of the DapL protein and identifying a correlation between total effect on protein dynamics *in silico* and inhibition measured *in vitro*.

## 2 ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. André Hudson for his steadfast support and guidance while I developed my thesis into a major multidisciplinary project. His expertise, assistance, and humor brought a much-needed feature to every day and has directly influenced my desire to continue my education. In addition, many thanks to the Hudson Lab research group for support, excitement, and friendship over the past 2.5 years, and thank you to Anthony Weatherhead from the University of Canterbury for his friendship, advice in creating a thesis, and determination to generate crystallography results. I would also like to thank my committee members, Dr. Gregory Babbitt and Dr. Michael Savka, for their constant support and troubleshooting guidance. Lastly, I thank and acknowledge the faculty and staff of the Thomas H. Gosnell School of Life Sciences (GSoLS) and the College of Science (COS) at the Rochester Institute of Technology (RIT) for the ongoing support of not only this Master's project, but also the Bachelor's in Biotechnology and Molecular Bioscience that brought me to this point. My success as a scientist is and always will be attributed to the support I received while at RIT. This work was also supported by a National Institutes of Health (NIH) award (R15GM120653).

3 TABLE OF CONTENTS

**1 ABSTRACT 3**

**2 ACKNOWLEDGEMENTS 4**

**3 TABLE OF CONTENTS 5**

**4 LIST OF FIGURES 7**

**5 LIST OF TABLES 9**

**6 BACKGROUND ON THE LYSINE BIOSYNTHETIC PATHWAY 10**

**7 CHAPTER 1: Developing a Model System in *Verrucomicrobium spinosum* 19**

*7.1 INTRODUCTION 19*

*7.2 METHODS 23*

7.2.1 Strains, Growth, and Transformation Conditions 23

7.2.2 Homologous Recombination 23

7.2.3 CRISPR/cas9 24

7.2.4 Growth Experiment 26

7.2.5 Scanning Electron Microscopy Imaging 26

*7.3 RESULTS 28*

*7.3.1* Phenotypic Analyses of *V. spinosum* 28

7.3.2 SEM Imaging 35

*7.4 DISCUSSION 44*

**8 CHAPTER 2: Comparative Molecular Dynamics Simulations of DapL 47**

*8.1 INTRODUCTION 47*

*8.2 METHODS 51*

8.2.1 Multiple Sequence Alignment 51

8.2.2 Structure Stabilization Testing 51

8.2.3 Compound Preparation 51

8.2.4 Molecular Docking 53

8.2.5 DROIDS Settings and Run Times 53

*8.3 RESULTS 53*

8.3.1 Multiple Sequence Alignment 53

8.3.2 Structure Stability Testing 55

8.3.3 Molecular Docking 56
8.3.4 Comparative Molecular Dynamics 62
8.3.5 DROIDS Summary 69
*8.4 DISCUSSION* *75*
**9 CONCLUSIONS** **79**
**10 SUPPLEMENTAL FIGURES** **81**
**11 REFERENCES** **86**

# 4 LIST OF FIGURES

Figure 1: The Lysine Biosynthesis Pathways 11

Figure 2: Peptidoglycan Structure and Organization 13

Figure 3: Model of V. spinosum DapL 14

Figure 4: cas9 targeting VsdapL 22

Figure 5: Recombination Schemes 24

Figure 6: gRNA Design 25

Figure 7: Transformation Outcomes 26

Figure 8: V. spinosum Growth Curve 28

Figure 9: VsDapL Complementation Assay 29

Figure 10: Examples of Unsuccessful Knockouts 30

Figure 11: gRNA quality and quantity as determined by NanoDrop 31

Figure 12: V. spinosum standard plating 32

Figure 13: Growth curve of cas9 knockouts 33

Figure 14: Kanamycin survival of cas9 knockouts 34

Figure 15: Two hours post transformation of wild type cells 35

Figure 16: Twenty-four hours post transformation of wild type cells 36

Figure 17: Two hours post transformation of mutant candidates from the cas9 complex and gene interruption cassette 37

Figure 18: Twenty-four hours post transformation of mutant candidates from the cas9 complex and gene interruption cassette 38

Figure 19: Two hours post transformation of mutant candidates from the cas9 complex and gene replacement cassette 40

Figure 20: Twenty-four hours post transformation of mutant candidates from the cas9 complex and gene replacement cassette 41

Figure 21: Twenty-four hours post transformation of mutant candidates from the cas9 complex 43

Figure 22: Antagonist Lead Compounds Identified for Use 52

Figure 23: VsDapL Multiple Sequence Alignment 54

Figure 24: VsDapL Dimer and Monomer Testing 55

Figure 25: Hydrazide Docking Structure 58

Figure 26: Rhodanine 59

Figure 27: Barbiturate 60

Figure 28: Thiobarbiturate 61

Figure 29: Binding Summary 62

Figure 30: DROIDS Plot and Images, VsDapL and Hydrazide 63

Figure 31: DROIDS Plot and Images, VsDapL and Rhodanine 65

Figure 32: DROIDS Plot and Images, VsDapL and Barbiturate 66

Figure 33: DROIDS Plot and Images, VsDapL and Thiobarbiturate 68

Figure 34: Atomic Fluctuation Summary 70

Figure 35: Significant fluctuation counts 71

Figure 36: Summary of Affected Active Site Residues 73

Figure 37: Rhodanine Double Docking 82

Figure 38: Full list of dampened and activated residues for each MD run 83

# 5 LIST OF TABLES

Table 1: Hydrazide Docking Statistics ..... 58

Table 2: Rhodanine Docking Statistics ..... 59

Table 3: Barbiturate Docking Statistics ..... 60

Table 4: Thiobarbiturate Docking Statistics ..... 61

Table 5: Summary of Active Site Changes in Fluctuation ..... 74

Table 6: RGI CARD Scan ..... 81

Table 7: VsDapL Inhibition ..... 85

# 6 BACKGROUND ON THE LYSINE BIOSYNTHETIC PATHWAY

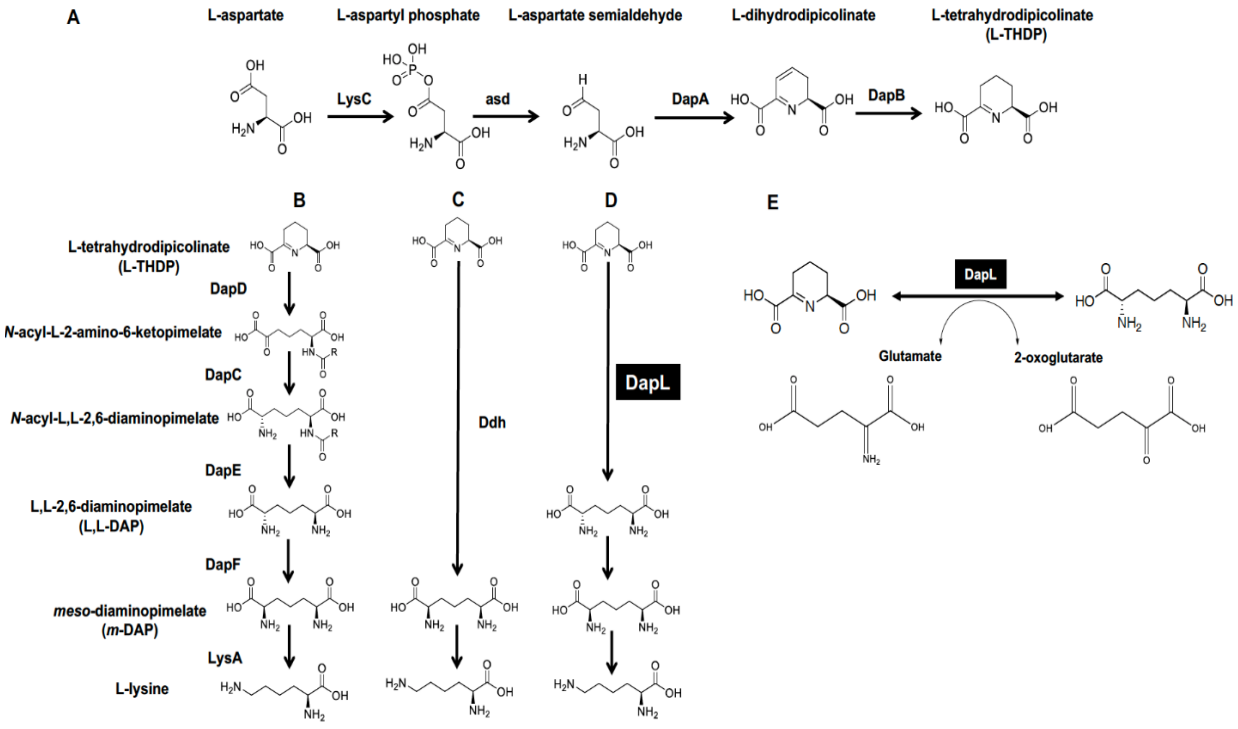
Recent years in medicine and technology have seen a marked increase in the incidence of multidrug resistant bacterial infections. Bacteria that cause these infections and develop such resistance, for example commonly among the *Pseudomonas*, *Chlamydia*, or *Staphylococcus* genera, can grow and propagate in the presence of some of the strongest antibiotic compounds currently known (Ventola, 2015). Antibiotic resistance can be acquired in different ways – some bacteria develop antibiotic resistance via accumulation of random mutations that allow it to propagate in the presence of an antibiotic while others develop resistance through transfer of external genetic factors (e.g. horizontal transfer of antibiotic resistance genes between organisms) (Munita & Arias, 2016). The rise of antibiotic resistant bacteria, even against antibiotics of last resort, highlights the need for novel targets in and approaches to antibiotic research and development to treat human disease.

Of the hundreds of amino acids found in nature, only twenty are ubiquitous in protein synthesis (twenty are deemed proteogenic). Of those twenty proteogenic amino acids nine are nutritionally essential for humans as they cannot be synthesized *in vivo* and must be obtained through diet. The pathways and proteins involved in the synthesis of these nine essential amino acids are attractive targets for antibiotic development because, in theory, targeting them would cause minimal toxicity in the patient other than their intended bactericidal effect.

Lysine is one of such essential amino acid, it cannot be synthesized by humans and the pathway is an attractive research target for antibiotic development. In organisms that have the ability, lysine biosynthesis occurs via one of two pathways: the alpha-aminoadipic pathway (AAA) is commonly found in yeast, fungi and some archaea while the diaminopimelate pathway (DAP) is employed by photosynthetic organisms and most prokaryotes (Nishida et al., 1999; Velasco,

Leguina, & Lazcano, 2002) and is the focus of this study. The DAP pathway has three main steps: (1) Conversion of aspartate to tetrahydrodipicolinate (THDP) by intermediary enzymes LysC, AspD, DapA, and DapB (*Figure 1A, abbreviations in Figure 1*), (2) conversion of THDP to *meso*-diaminopimelate (*m*-DAP), and (3) conversion of *m*-DAP to Lysine by LysA (*Figure 1*). There are four known variants of the DAP pathway, distinguished by differences in *m*-DAP generation: the acyl-diaminopimelate pathways (*Figure 1B*), the *meso*-diaminopimelate dehydrogenase (Ddh) pathway (*Figure 1C*), and the recently identified l,l-diaminopimelate aminotransferase (DapL) pathway (Hudson, Singh, Leustek, & Gilvarg, 2006) (*Figure 1D*).

*Figure 1: The Lysine Biosynthesis Pathways*




*(A) The synthesis of THDP from aspartate, (B) The acyl pathway, (C) the Ddh pathway, (D) the DapL pathway, and (E) the reversible transamination reaction catalyzed by DapL.*

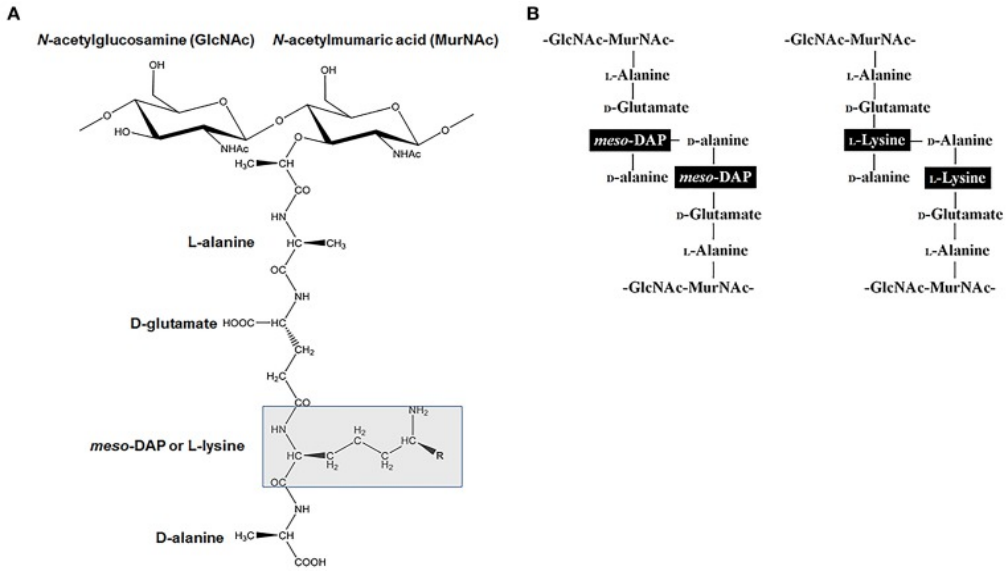
*Abbreviations: DapC, acyl-diaminopimelate aminotransferase; DapE, acyl-diaminopimelate deacylase; DapD, 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate-N-acyl-transferase; DapF, acyl-diaminopimelate deacylase, DapL, L,L-diaminopimelate aminotransferase; Ddh, diaminopimelate dehydrogenase; LysA, diaminopimelate decarboxylase. Adapted from (Triassi et al., 2014).*

In the acyl-diaminopimelate pathways, most commonly found in bacterial species, DapD and DapC enzymes produce acetylated (via acetyl-CoA) or succinylated (via succinyl-CoA) intermediates from THDPA, followed by production of L,L-DAP by DapE and its conversion to *m*-DAP by DapF. The acyl pathways constitute two of the four DAP pathway variants. The third variant of the pathway, the Ddh pathway converts THDPA directly to *m*-DAP in a single reaction by Ddh. Lastly, the DapL pathway, identified in plants and some bacterial species and first identified in *Arabidopsis thaliana*, utilizes DapL in place of the DapC, DapD, DapE, DapF, or Ddh enzymes found in other organisms in the conversion of THDP to *m*-DAP. It converts THDP to L,L-DAP via a single, direct, reversible transamination reaction after which the DapF enzyme converts L,L-DAP to *m*-DAP for further use (Hudson et al., 2005, 2006) (*Figure 1E*).

The lysine biosynthetic pathway is also a potential antibiotic target because of its involvement in downstream pathways, specifically in cell wall construction in most bacteria. In Gram-negative bacteria, *m*-DAP is in the third position of a peptide stem that links layers of *N*-acetylglucosamine – *N*-acetylmumaric acid disaccharides to form the peptidoglycan cell wall; in Gram-positive bacteria, lysine is in the center of the linking peptide chain between layers of disaccharides (Hutton, Perugini, & Gerrard, 2007) (*Figure 2*). Without the cross-linking of peptidoglycan, the cell wall loses stability and is unable to adequately protect the cell and maintain separation of the internal and external environment. Without this stability afforded by

the cell wall, the bacterium will lose stability and eventually rupture. Because of its critical role in peptidoglycan crosslinking and protein synthesis as well as its lack of presence in mammals (specifically humans), the lysine biosynthetic pathway is an attractive target for the development of novel antibiotics with minimal predicted side effects.

*Figure 2: Peptidoglycan Structure and Organization*

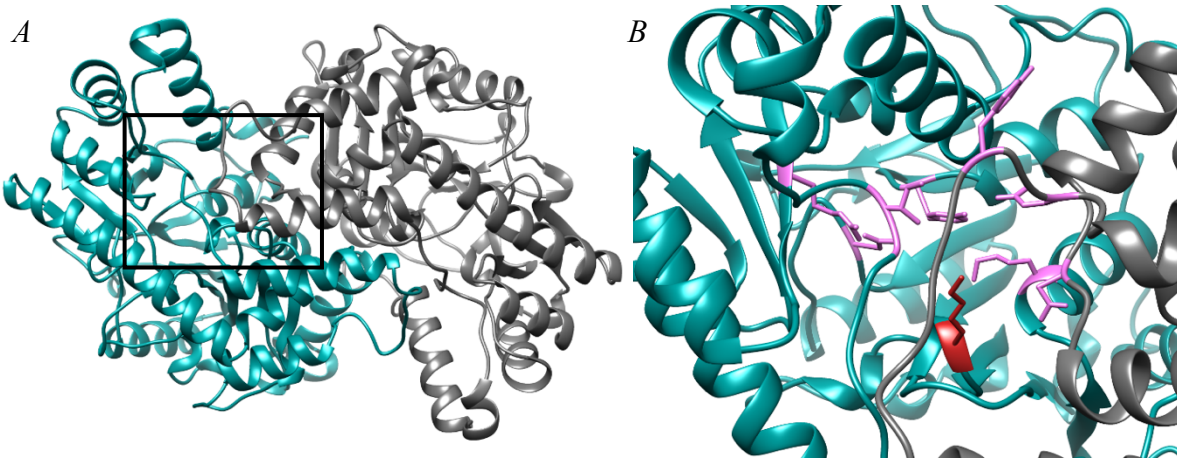



*(A) Structure of the monomeric subunit of peptidoglycan, with the cross-linking highlighted in the grey box. (B) Cross-linking by meso-DAP (Gram negative) and Lysine (Gram positive) bacteria. Adapted from (Triassi et al., 2014).*

The key enzyme in the DapL pathway, L,l-diaminopimelate aminotransferase (DapL) is a homodimer whose two subunits are each ~40 kD in size (*Figure 4A*) (Biasini et al., 2014). The enzyme in its native form contains two active sites and functions on a hinge; only one site is active at a time due to a conformational change in the other site when bound with a substrate. Each subunit is composed of a major and minor arm, the major arm composing the dimer interface for each subunit. The DapL enzyme, via a classic bimolecular ping-pong mechanism

(Velick & Vavra, 1962), transfers an amino group plus a proton and electron pair from a donor molecule to an acceptor molecule. In aminotransferases, this process occurs in two steps. First, the amino group from the donor molecule forms a Schiff based structure with a pyridoxal-5'-phosphate (PLP) molecule that is covalently bonded to the conserved lysine residue in the active site (*Figure 3)* followed by the transfer of the amino group to the acceptor molecule (Liepman & Olsen, 2004). As previously mentioned, the DapL enzyme catalyzes a reversible transamination reaction; a glutamate acts as the donor molecule and THDP acts as the acceptor in the forward reaction while L,L-DAP acts as the donor molecule and 2-oxoglutarate acts as the acceptor in the reverse reaction (Hudson et al., 2006).

*Figure 3: Model of V. spinosum DapL*




*The Verrucomicrobium spinosum DapL protein (VsDapL), modeled by SWISS-MODEL, is a homodimer with two active sites. (A) The full enzyme colored by subunit, one active site highlighted by a black box. (B) The active site close view with key residues (Y74, E77, and N294 on the chain opposite the pocket and Y134, I43, G44, R390, N189, and K111) colored in pink and the conserved K251 that binds with PLP (pyridoxal-5'-phosphate) colored in red.*

As previously mentioned, distribution of the lysine biosynthetic pathway variants is widespread across organisms. The alpha-aminoadipic pathway (AAA) is commonly found in yeast, fungi and Archaea while the diaminopimelate pathway (DAP) is employed by photosynthetic organisms and most prokaryotes (Nishida et al., 1999; Velasco et al., 2002). In addition, most bacterial species employ the acyl-DAP pathway, a handful of bacterial species employ the Ddh pathway, and plants and some other bacterial species employ the DapL pathway (Nishida et al., 1999; Velasco et al., 2002). The DapL pathway specifically has been identified in the genomes of 13% of all bacteria (Triassi et al., 2014), a narrow distribution containing bacterial species of medical and biotechnological relevance, plants, and some archaea.

Among the bacterial species employing the DapL pathway, those belonging to the phyla *Chlamydophila, Treponema, Leptospira,* and *Bacteroides* are especially of interest. Each contain organisms of extreme relevance and understanding the critical metabolic pathways of those species would provide beneficial insight. For example, *Chlamydia trachomatis* is a pathogen and the causative agent of one of the most reported sexually transmitted diseases in the United States and the leading cause of infection-mediated blindness worldwide (Mishori, McClaskey, & WinklerPrins, 2012). It is a Gram negative, obligate intracellular bacterium with high rates of lateral gene transfer between serotypes and largely uncharacterized genetic properties (Mabey, 2008). *Treponema pallidum* is the causative agent of human Syphilis, an infection that has maintained its status as a significant cause of human disease regardless of the availability of an effective treatment (Burstain, Grimprel, Lukehart, Norgard, & Radolf, 1991). *T. pallidum* is also commonly found as a secondary infection in HIV-immunocompromised patients, and complications from general infection include neurosyphilis and retinitis (Berry, Hooton, Collier, & Lukehart, 1987). Infection with *Leptospira interrogans* causes leptospirosis, an illness with a

wide range of symptoms presenting with flu-like symptoms and occasionally progressing to a hemorrhagic fever (Ansdell, 2012; Lindsay, Bone, & Fuller, 2010). In severe cases it is treated with intravenous antibiotics and kidney and liver dialysis treatments. *Bacteroides fragilis* is commonly found as a microbiotic component of the human gastrointestinal tract, but causes significant anaerobic infections upon a breach of that system including infections of the head, neck, pelvis, soft tissues, and nervous system (Finegold & Sussman, 2002). In addition, enterotoxigenic *B. fragilis* causes significant gastrointestinal illness (Wexler, 2007). Each of the aforementioned species are the causative agents of significant and serious human disease, and each of the aforementioned species either employ the DapL pathway or have a close relative in the same genus that employs the DapL pathway for lysine biosynthesis.

Also contained in the distribution of the DapL pathway is the PVC superphylum, a monophyletic group consistently recovered in 16S rRNA trees containing phyla *Planctomyces, Verrucomicrobia, Chlamydiae, Lentisphaerae,* and candidates *Poribacteria* and OP3 that span interesting groups. *Planctomyces* are commonly found in aquatic and terrestrial ecosystems (Dedysh, Pankratov, Belova, Kulichevskaya, & Liesack, 2006; Vergin et al., 1998). Species from this phylum have reported properties for useful biotechnological application in numerous areas, ranging from formaldehyde detoxification to denitrification and the removal of ammonia from water sources (Wagner & Horn, 2006). *Chlamydiae*, members of whom were previously mentioned, is commonly identified as containing extremely successful human pathogens, but it also contains some commensal and environmental species. *Verrucomicrobia* contains widely abundant environmental microorganisms and is suspected to comprise up to 10% of the total bacteria in soil (Sangwan, Chen, Hugenholtz, & Janssen, 2004). Though *Verrucomicrobia* are not known to be pathogenic, several species live in association with other eukaryotes. There is

also evidence that *Verrucomicrobium spinosum* utilizes the non-flagellar type III secretion system like some chlamydial species, a secretion system known to mediate interaction between the bacteria that contain it and the target eukaryotic cells (Pallen, Beatson, & Bailey, 2005). Interestingly, numerous eukaryotic genes have been identified in the genomes of these organisms, indicating a high rate of lateral gene transfer between these bacterial species and their eukaryotic host cells. Though species of *Verrucomicrobia* have been found independent of any eukaryotic cells, no free-living species of *Chlamydia* or *Poribacteria* have been found. The PVC superphylum contains a group of bacterial species with interesting properties as well as highly relevant bacterial species that should be isolated and investigated further for biotechnological and clinical application.

Though its distribution is narrow, the DapL pathway is employed by species with unique properties and those of biotechnological and clinical importance. A large majority of species in the PVC superphylum employ an anaerobic metabolism, are obligate intracellular organisms, or are pathogenic to humans (e.g. *Chlamydia trachomatis*). As such, manipulation of most organisms that employ the DapL pathway is a formidable task. However, *Verrucomicrobium spinosum* is a phenotypic outlier. It is a slow-growing, heterotrophic bacterium that is non-pathogenic to humans and the closest free-living relative of *C. trachomatis* based on examination of both widely distributed proteins and shared inserts (Griffiths & Gupta, 2007). Evidence of peptidoglycan in the cell wall of *Chlamydia* was also recently discovered and validated (Packiam, Weinrick, Jacobs, & Maurelli, 2015; Pilhofer et al., 2013), further supporting the use of *V. spinosum* in the development of a model system to evaluate the DapL pathway for the development of narrow-spectrum therapeutics for bacterial infections.

We hypothesized that inhibition of L,L-diaminopimelate aminotransferase will cause a bactericidal effect through prevention of peptidoglycan crosslinking and protein synthesis because of the involvement of both *meso*-diaminopimelate/lysine in the peptidoglycan crosslinking of the bacterial cell wall and the status of lysine as a proteogenic amino acid. In this thesis, the essentiality of the gene encoding the DapL protein in *Verrucomicrobium spinosum* was assessed as a step toward the development of a model system and the dynamic behavior of the enzyme in its native state was compared to its behavior when complexed with antagonistic lead compounds in comprehensive, comparative molecular dynamics simulations.

# 7 CHAPTER 1: Developing a Model System in *Verrucomicrobium spinosum*

## 7.1 INTRODUCTION

DapL is a good target for narrow spectrum antibiotic development because of both its limited distribution in prokaryotes and its dual involvement in lysine biosynthesis and peptidoglycan crosslinking of the cell wall. The DapL pathway was identified in approximately 13% of the genomes of all bacteria and archaea, including those of interest detailed previously. DapL has been proven essential in plants that contain the pathway; in instances where the gene is eliminated or inactivated the plant cannot survive without 2,6-diaminopimelate (DAP) supplementation (Dobson, Girón, & Hudson, 2011). While the same has been assumed true for bacteria that contain the DapL pathway, it has never experimentally investigated and ambiguity surrounding both lysine biosynthesis and aminotransferases in bacteria also underscores the need to direct investigation. For example, evidence of two pathways for lysine biosynthesis has been found in two bacterial species; *B. fragilis* and *C. thermocellum* contain both the DapL and Ddh pathways (Hudson et al., 2011). However, no other bacterial species to date have been identified as containing dual pathways. In addition, some aminotransferases are proven to be promiscuous, they can act on multiple substrates. It is currently not known if the same may occur in bacteria that employ the DapL pathway as the sole producer of lysine, or if it is fully dependent on a functional *dapL*. The development of a model system that can be cultured and is non-pathogenic is an ideal approach to study the essentiality of the pathway and develop inhibitors of the DapL enzyme.

As the chosen bacterium in this project, *Verrucomicrobium spinosum* employs the DapL pathway as its sole producer of lysine and *m*-DAP for downstream peptidoglycan crosslinking (Nachar et al., 2012). It is a Gram-negative, heterotrophic, slow growing organism commonly

found in freshwater environments. While it is known to associate with eukaryotic hosts, *V. spinosum* is non-pathogenic to humans. This organism was also chosen because *Verrucomicrobium* is a member of the PVC superphylum (Wagner & Horn, 2006), composed of phyla *Planctomycetes, Verrucomicrobia*, *Chlamydiales,* and *Lentisphaerae,* of whom are host to medically and biotechnologically important species.

Genetic manipulation of *V. spinosum* has also been investigated in one laboratory; it was previously manipulated via random transposon insertion (Domman, Steven, & Ward, 2011) with sufficient recovery and antibiotic resistance. In addition, the presence of recombination proteins A (RecA) and R (RecR), involved in replication-associated and homologous recombination, respectively in the annotated genome, and the suspected high lateral gene transfer rate between it and other organisms suggest success in further genetic manipulation. However, the whole genome sequence indicates no evidence of endogenous plasmids, and there has been no data to support the replication and maintenance of a broad host-range plasmid in this organism. The presence of genetic recombination enzymes and successful Tn5 mutagenesis with *V. spinosum* indicate that, while many properties of the organism are largely unknown, it is a good candidate for genetic knockout to test essentiality, and for further genetic and phenotypic studies (Nachar et al., 2012).

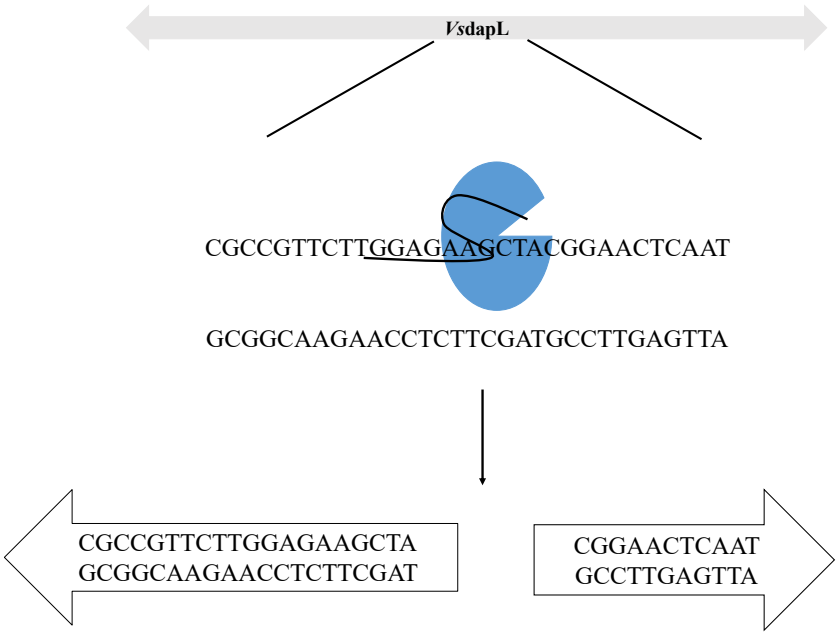
One of such techniques for example, homologous recombination is a type of biological genetic recombination where two independent nucleotide sequences are exchanged, guided by base-pairing interactions of the homologous nucleotide sequences. This event is mediated by recombination enzymes. The DNA double helix separates at the homologous site and one strand creates a heteroduplex joint with the matching strand on the other DNA molecule. After cleavage and rejoining, a portion of one DNA molecule is swapped with the other molecule (Alberts et al.,

2002). Its application in gene editing simply involves the use of a desired insert of choice (e.g. a selectable antibiotic resistance marker nucleotide sequence) flanked by sequences homologous to the targeted insertion site (e.g. *dapL*). Transformation and selection for the desired phenotype is expected to result in a mutated, non-functional *dapL*.

Another genetic editing tool, CRISPR (Clustered Regularly InterSpaced Palindromic Repeats) has become very popular in recent years. First identified as a proteolytic enzyme and a component of bacterial immune defense mechanisms, CRISPR was characterized as a region of the bacterial genome where the 3' end was marked with direct repeats spaced with a 32-nucleotide spacer and a CRISPR-associated (*cas*) gene directly next to the repeating region (Ishino, Shinagawa, Makino, Amemura, & Nakata, 1987). The spacer sequence was identified as a sequence homologous to a region targeted by Cas (Garneau et al., 2010; Jansen, Embden, Gaastra, & Schouls, 2002). The CRISPR/Cas system is fully self-contained and can be expressed in a vector with only the direct repeats, *cas*, and targeting sequence (Sapranauskas et al., 2011). The most common current tool is the Type II CRISPR/*cas9* locus. There are three essential regions in the genetic makeup of this tool: (1) a *tracrRNA* gene, (2) the *cas* gene, and (3) the CRIPSR repeat-spacer array. The CRISPR repeat-spacer array is transcribed into pre-crRNA, the *cas* gene is transcribed and translated into the Cas protein (in this case Cas9), and the *tracrRNA* gene is transcribed into tracrRNA. Post-transcription processing by *cas9*, tracrRNA, and an RNase modifies the pre-crRNA to form a set of unique tracrRNA/*cas9*/crRNA complexes that then cleave target DNA sequences based on an alignment with the tracrRNA sequence (Chylinski, Le Rhun, & Charpentier, 2013) (*Figure 4*). The tracrRNA/crRNA complex is sometimes replaced with a single gRNA (the guide RNA) for ease of build and distribution. This guide RNA contains the targeting sequence (like the tracrRNA) as well as the sequences

necessary for interaction with the *cas* protein (Mali et al., 2013). Target genes are modified by either random insertion due to DNA breakage or by introducing a DNA sequence into the cell for incorporation after DNA breakage and the activation of DNA native repair mechanisms.

*Figure 4: cas9 targeting VsdapL*




*cas9 introducing a double-stranded break in target DNA, as guided by the gRNA upon association with cas9.*

Here, methods of generating a *V. spinosum* mutant deficient in *dapL* are reported, with the purpose of providing data that supports the essentiality of the gene in bacteria that employ the DapL pathway as the sole producer of lysine and peptidoglycan crosslinking.

## 7.2 METHODS

### 7.2.1 Strains, Growth, and Transformation Conditions

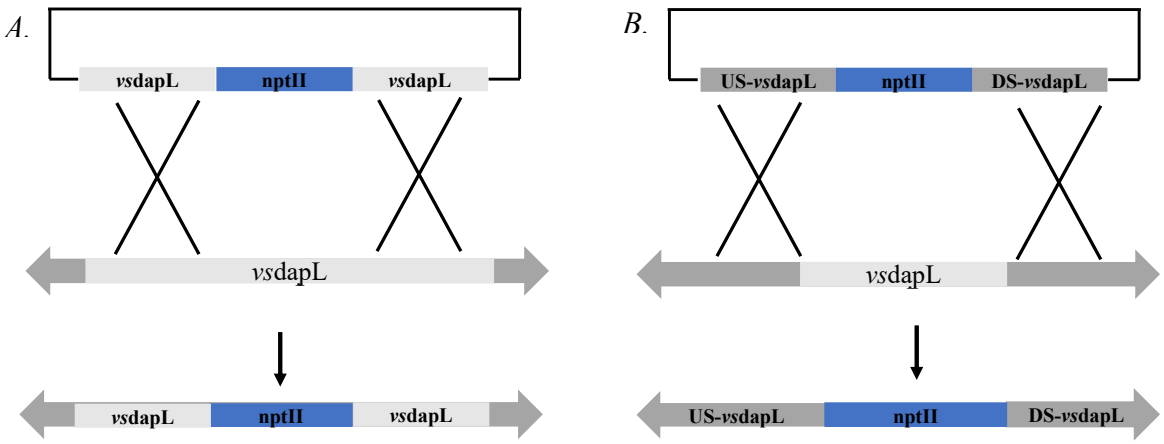
*V. spinosum* DSM 4136$^{T}$, obtained from the American Type Culture Collection (ATCC #43997) was cultured on M13 minimal media. For all transformations, a stationary culture (measured at $OD_{600}$ 0.3) was sub-cultured in 5 ml and grown to exponential phase at an $OD_{600}$ 0.15 (about 30 hours) before electrocompetent cell preparation. The cells were pelleted, washed twice in cold sterile $dH_20$ and twice in cold sterile 10% glycerol before resuspension in ~100 ul cold sterile 10% glycerol. Electroporation occurred with two pulses at 1.0 kV, followed by resuspension in equal volume (typically 5 ml) 200 ug/ml DAP-supplemented M13 and recovery of 2 hours and 24 hours before plating of 100 ul on M13 agar supplemented with Dap and antibiotic. Plates were incubated for ~one week at 30°C in a humid environment for colony formation.

### 7.2.2 Homologous Recombination

Two genetic knockout approaches were tested: gene interruption and gene replacement with a selectable marker (neomycin phosphatase NptII, a kanamycin resistance gene) (*Figure 5*). To introduce a gene interruption, *nptII* was cloned, in frame, in the center of *dapL* to generate a knockout cassette. To introduce a full gene replacement, *nptII* was closed directly between *dapL* upstream and *dapL* downstream flanking regions to generate a knockout cassette. The knockout cassette was then cloned into replication-deficient pUC57 (lacking the proper origin of replication) vector before transformation into *V. spinosum*. The vectors were constructed by Bio Basic Inc. custom ordering services (https://www.biobasic.com/). Both knockout cassettes were also PCR amplified and purified for linear transformation into *V. spinosum.* To inhibit the native restriction endonucleases, an inhibitor was also transformed into *V. spinosum* during each experiment ("TypeOne$^{TM}$ Restriction Inhibitor," n.d.). After transformation and recovery, cells

were plated on media supplemented with DAP and Kanamycin for nutrient supplementation and antibiotic selection, respectively.

*Figure 5: Recombination Schemes*



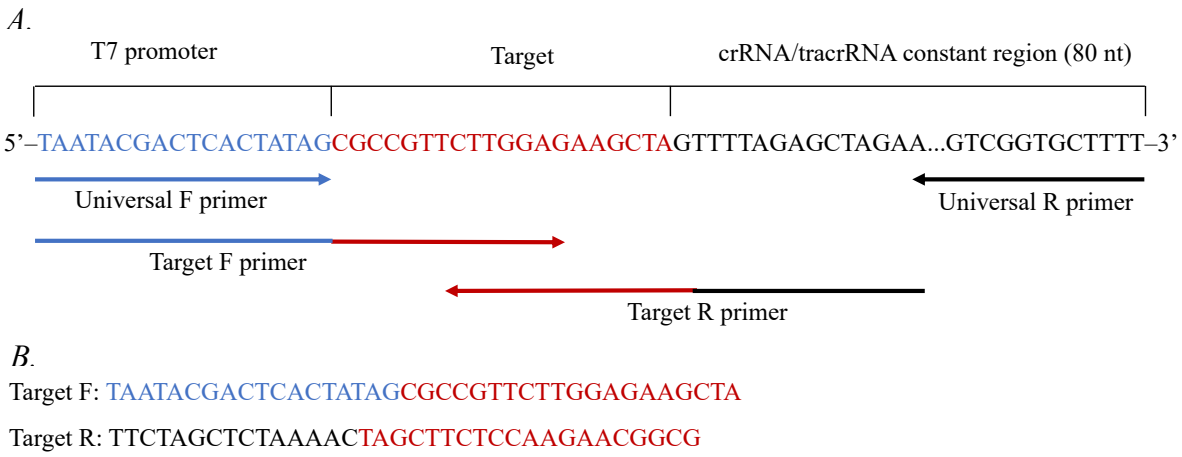

*(A) Illustrated knockout via homologous recombination (double X's) mediated by a gene interruption cassette containing sequences homologous to the VsdapL sequence and (B) illustrated knockout via homologous recombination mediated by a gene replacement cassette containing sequences homologous to regions directly upstream ('US') and downstream ('DS') of the VsdapL sequence.*

### 7.2.3 CRISPR/cas9

The first implementation of the CRISPR/cas9 system in *V. spinosum* used the pCas9 plasmid (Jiang, Bikard, Cox, Zhang, & Marraffini, 2013). A target sequence was constructed in the form of forward and reverse-complementary oligos, annealed together, and digested with BsaI. The digested sequence was ligated into the BsaI site of pCas9 between the direct repeating segments. Clones were screened for proper insert sequence via restriction digest and Sanger sequencing before transformation and expression testing in *V. spinosum*.

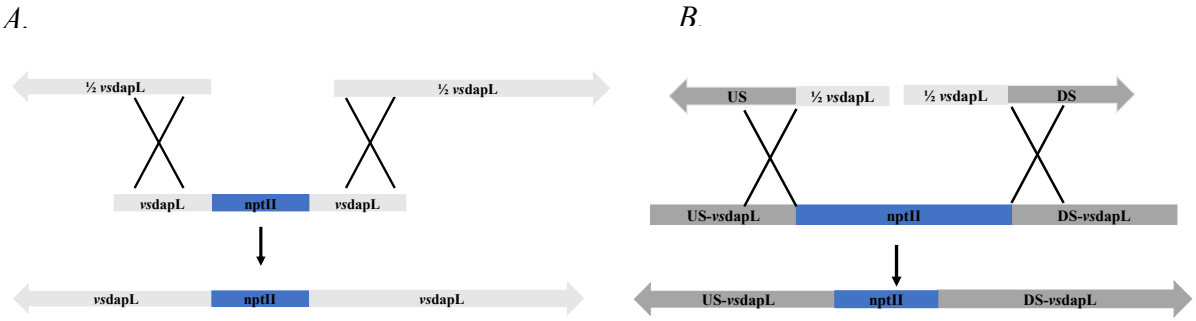
The second implementation of the CRISPR/cas9 system was an *in vitro* experiment utilizing a gRNA synthesis protocol and transfection-ready Cas9 from ThermoFischer (Liang, Potter, Kumar, Ravinder, & Chesnut, 2017). A target sequence in DapL was identified and primers were synthesized to include a T7 promoter and the beginning of the target sequence (F primer) as well as a tracrRNA constant region corresponding to Cas9 (R primer). gRNA design to target *V. spinosum dapL* is depicted in *Figure 6*. After PCR synthesis of the gRNA template DNA, *in vitro* transcription and RNA purification was performed for high quality, high concentration gRNA. The gRNA was briefly incubated with Cas9 in a ratio of 1.2:1 gRNA:Cas9 for complex association prior to transformation into *V. spinosum* with either gene interruption cassettes, gene knockout cassettes, or no addition (*Figure 7*). After transformation, $OD_{600}$ was tracked during recovery and before plating.

*Figure 6: gRNA Design*




*The design of guideRNA targeting the VsdapL gene containing (A) a T7 promoter region (blue), target sequence (red), and crRNA constant region (black) with primer alignment below. (B) Target forward and reverse primers designed to amplify the constant region (black), target sequence (red), and T7 promoter (blue).*

*Figure 7: Transformation Outcomes*




*(A) Illustrated knockout via homologous recombination (double X's) mediated by a gene interruption cassette containing sequences homologous to the vsdapL sequence after double stranded breakage mediated by Cas9 and (B) illustrated knockout via homologous recombination mediated by a gene replacement cassette containing sequences homologous to regions directly upstream ('US') and downstream ('DS') of the vsdapL sequence after double stranded breakage mediated by Cas9.*

### 7.2.4 Growth Experiment

Post-transformation cells were resuspended in equal volume (to pre-transformation preparation) M13 or DAP-supplemented M13 and their growth was tracked over 48 hours. The cells from each transformation condition, after 48 hours recovery in 200 ug/ml DAP-supplemented M13 were pelleted, washed twice in M13, and resuspended in 50 ug//ml kanamycin-supplemented M13 or DAP and kanamycin-supplemented M13 for growth tracking.

### 7.2.5 Scanning Electron Microscopy Imaging

The cell fixing protocol was obtained from Dr. Anutthaman Parthasarathy at RIT. 1.5 ml cells from each culture were reserved after two and twenty-four hours of recovery, pelleted, and stored in ~10 ul M13 media at 4° C until fixing. The pellet was resuspended, the entire contents were

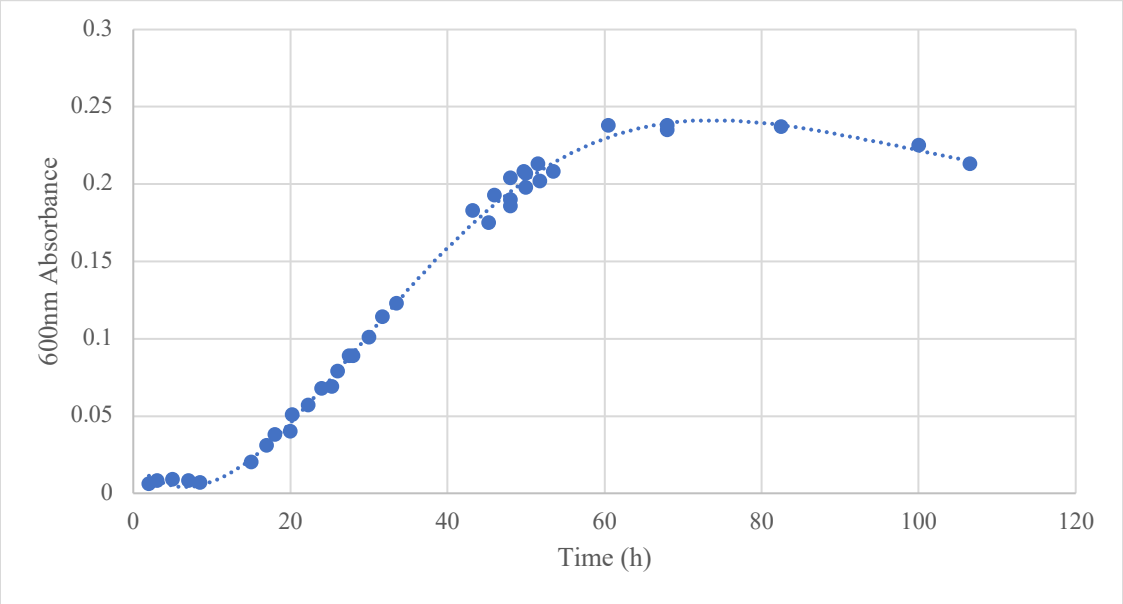
pipetted onto a glass slide cover and soaked in 2% glutaraldehyde solution for two hours. Excess solution was decanted off after the culmination of two hours and dehydrated in a graded ethanol series with ten-minute washes of 50%, 70%, 80%, 95%, and 100% ethanol solutions. Samples were stored with parafilm at room temperature until imaging at 10kV scanning electron microscopy.

## 7.3 RESULTS

### *7.3.1* Phenotypic Analyses of *V. spinosum*

*V. spinosum* has a slow growth rate spanning several days and density-limited growth in liquid culture (*Figure 8*). It has a doublimg time of about 20 hours. The growth curve indicates optimal timing for genetic manipulation occurs between 12 and 48 hours post-inoculation, during its exponential growth phase and at an $OD_{600}$ between 0.05 and 0.20. For adequate growth on solid media, *V. spinosum* also needs optimal humidity, minimal media, and a low cell density if plating from a liquid culture.

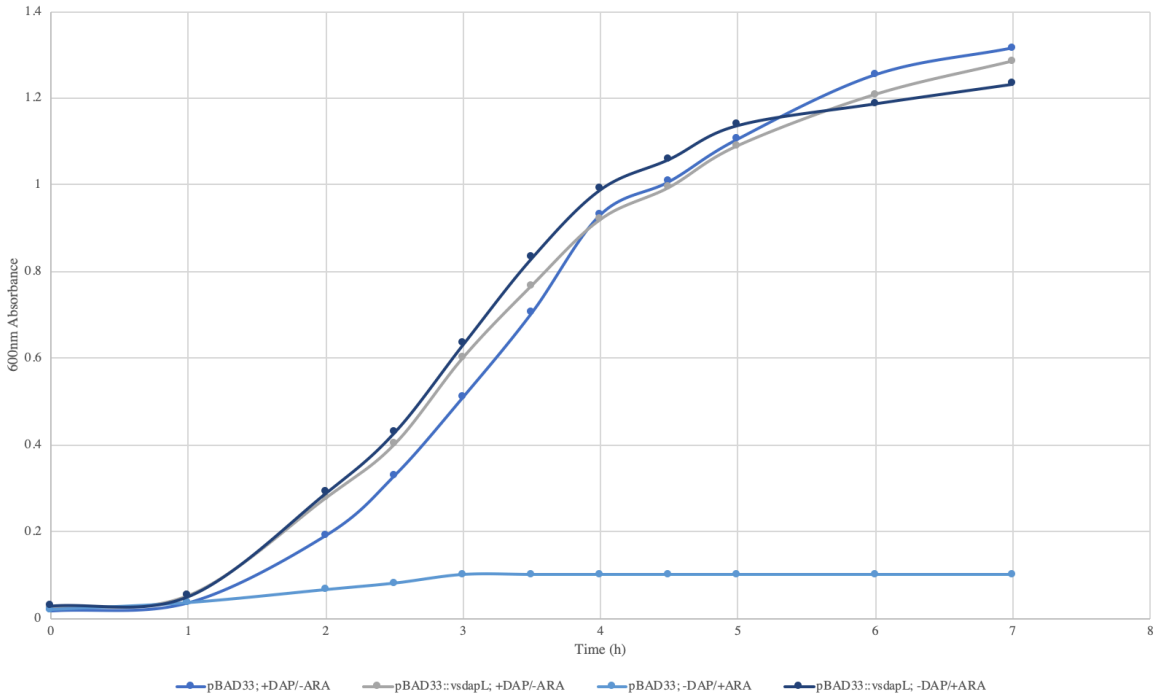
*Figure 8: V. spinosum Growth Curve*




*Absorbance at 600 nm of V. spinosum liquid culture over time at 30°C. It has a relatively short lag phase before entering exponential growth for a period between 10 and 60 hours after inoculation until an $OD_{600}$ of 0.25, after which point cell death starts to prevail.*

*VsdapL*, previously cloned into pBAD33 in the Hudson lab, successfully complements *E. coli* mutants deficient in the Dap D/E enzymes of the lysine biosynthetic pathway (*Figure 9*). Mutants complemented with the VsdapL construct grew at the same rate as mutants without *dapL* supplemented with DAP.

*Figure 9: VsDapL Complementation Assay*

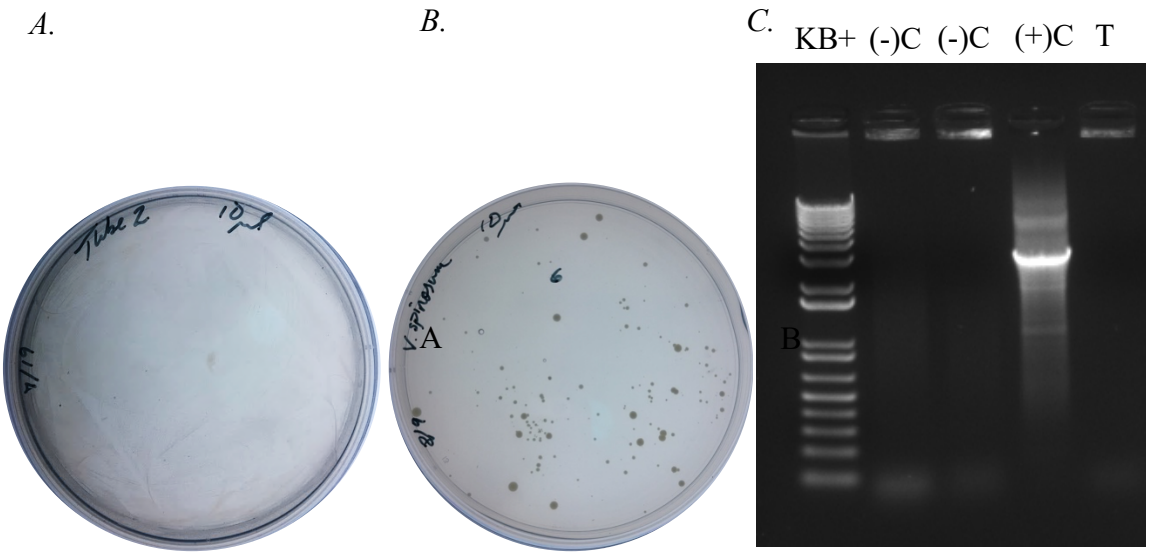



*Absorbance at 600 nm of cultures of DapD/E deficient E. coli mutants supplemented with DAP, complemented with the VsDapL enzyme, and without any supplements. Optical density measurements were taken over the course of seven hours.*

However, no knockout attempts in *V. spinosum* to investigate *dapL* requirement via homologous recombination were successful. Of the forty experiments performed testing homologous recombination as a knockout tool, fifteen resulted in no growth upon plating, ten resulted in lawns or smearing on plates because of bacterial over-saturation, and any remaining that were

recoverable lacked the target genotype (evidence of the kanamycin resistance gene or the *dapL* via confirmatory PCR). Most genotype screens resulted in proper controls with no result in the test case (*Figure 10*).

*Figure 10: Examples of Unsuccessful Knockouts*

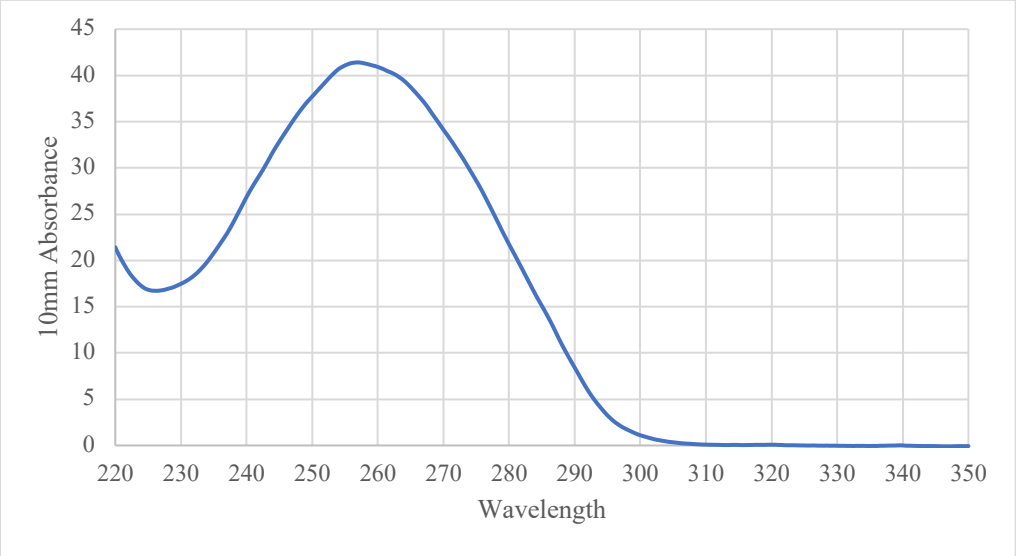



*Images of M13 agar plates from homologous recombination experiments that show (A) lawn/smearing of cells due to culture oversaturation (B) recoverable cell phenotypes for screening and (C) genetic screens via PCR that show a lack of target genotype.*

No recoverable phenotypes were propagated from the experiments with pCas9. Cells that showed potential phenotype lacked the presence of any genotype indicating either plasmid maintenance or knockout. In addition, the *V. spinosum* genome contains evidence of a potential chloramphenicol acetyltransferase gene (*Table 6*), which may confer chloramphenicol resistance and thus eliminate pCas9 dependence under chloramphenicol selection.

A high concentration of extremely pure RNA was generated for use with the platinum cas9 nuclease in the *in vitro* experiments (*Figure 11*), with a working stock concentration of ~1600 ng/ul. Subsequent complex formation with cas9 and transformation into *V. spinosum*, however, resulted in no recoverable target phenotypes. Plating was standardized to yield single colonies in the controls (*Figure 12*), but no other growth was observed in the test conditions.

*Figure 11: gRNA quality and quantity as determined by NanoDrop*



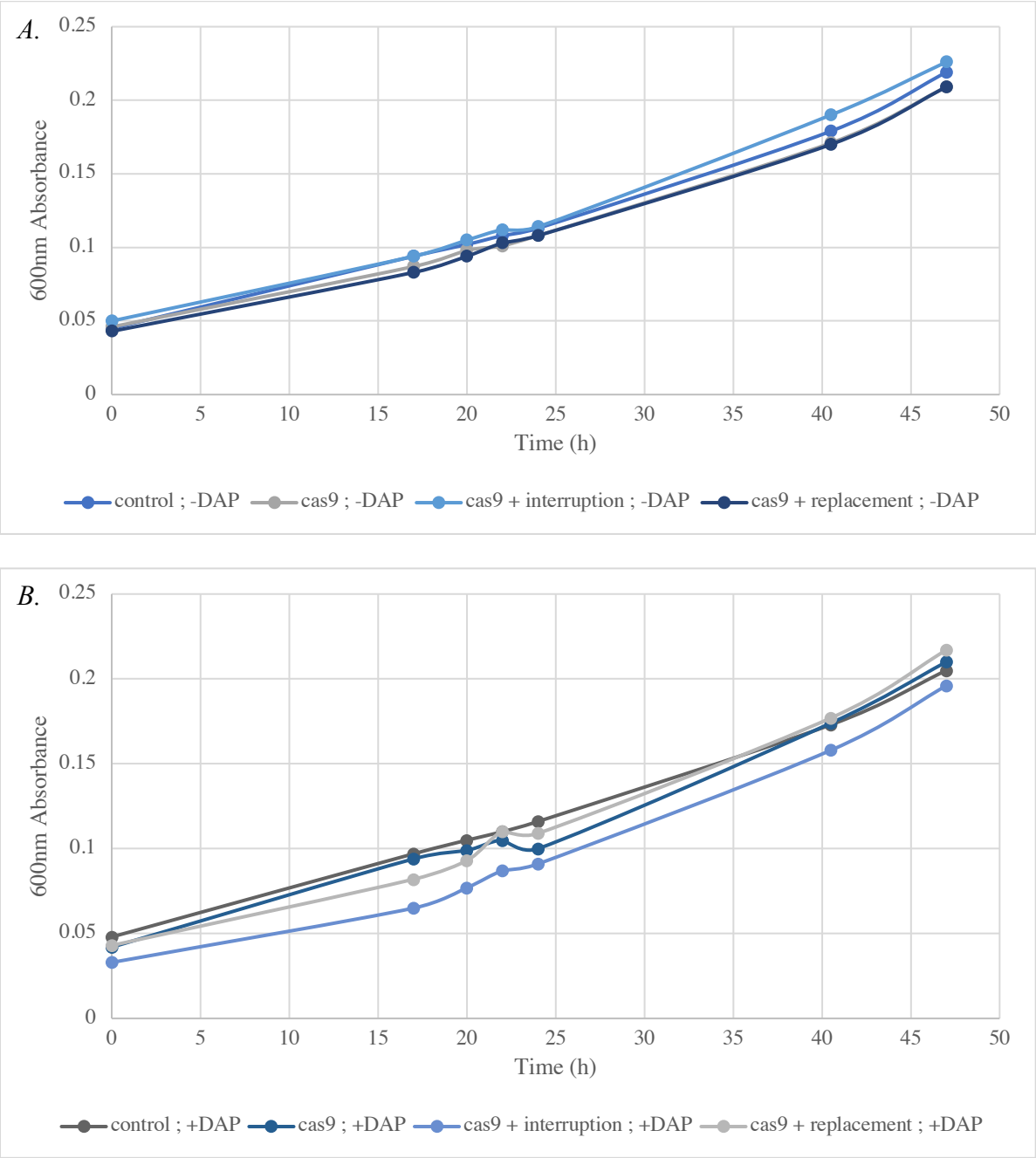

*Absorbance at 10nm of synthesized gRNA samples, measured by Nanodrop. The concentration was measured at 1636.52 ng/ul, with an A260/A280 of 1.881 and A260/A230 of 2.342.*

*Figure 12: V. spinosum standard plating*

*Wild-type V. spinosum plated on M13 for single colonies from a liquid culture after a transformation experiment.*

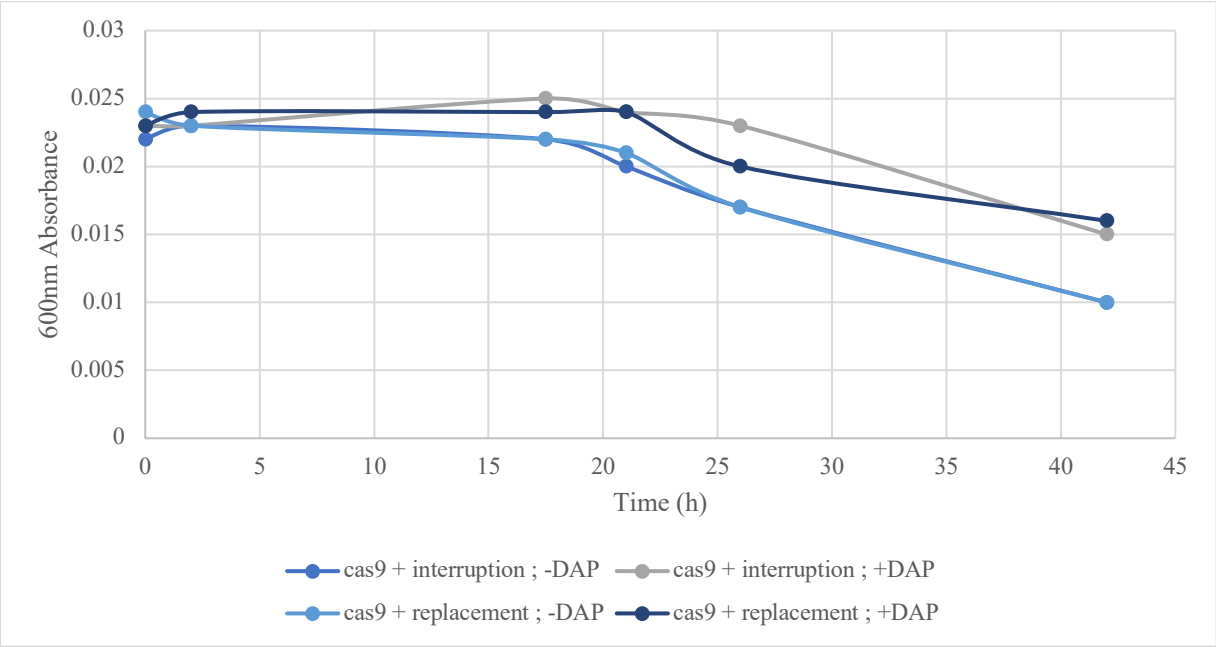
In liquid culture after the Cas9 transformation experiments, *V. spinosum* transformants grew both without and with the presence of DAP (*Figure 13*). The growth trajectories differ slightly, but there was minimal or no cell death or lag phase of growth. Upon introduction to kanamycin, cells transformed with the knockout cassettes died slightly slower than the controls (*Figure 14*).

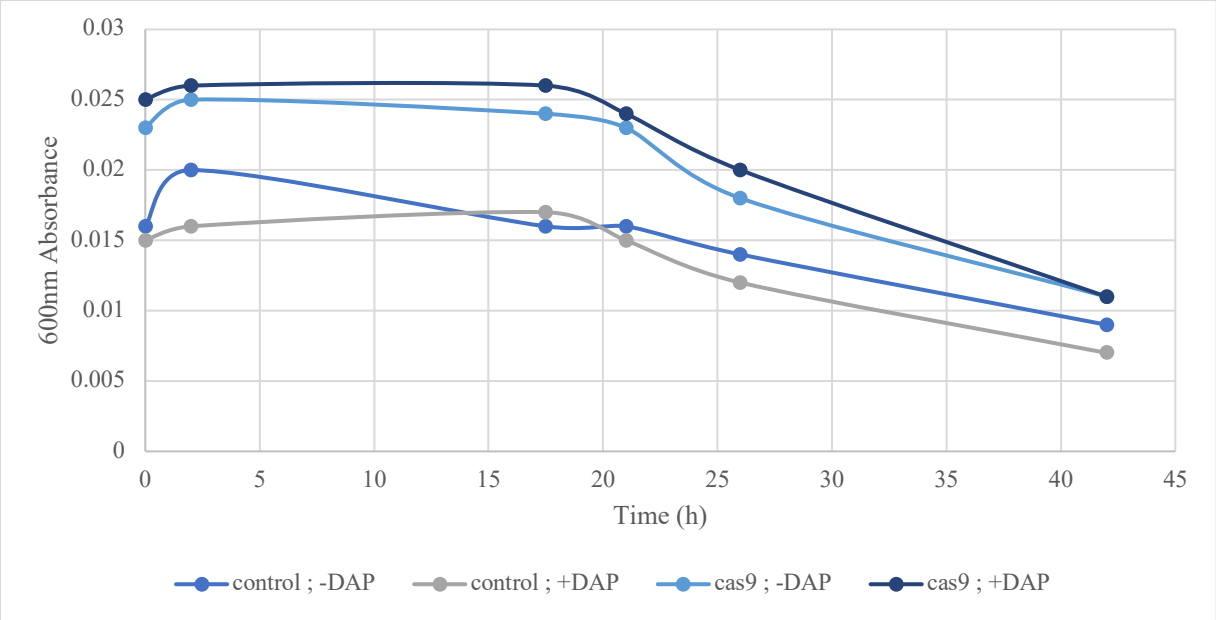
*Figure 13: Growth curve of cas9 knockouts*




*Absorbance at 600 nm of mutant and wild type V. spinosum. Samples included cultures transformed with a cas9/guideRNA complex, the complex combined with a gene replacement or gene interruption cassette, or no transformed sample cultured in either (A) M13 or (B) M13 supplemented with DAP.*

*Figure 14: Kanamycin survival of cas9 knockouts*



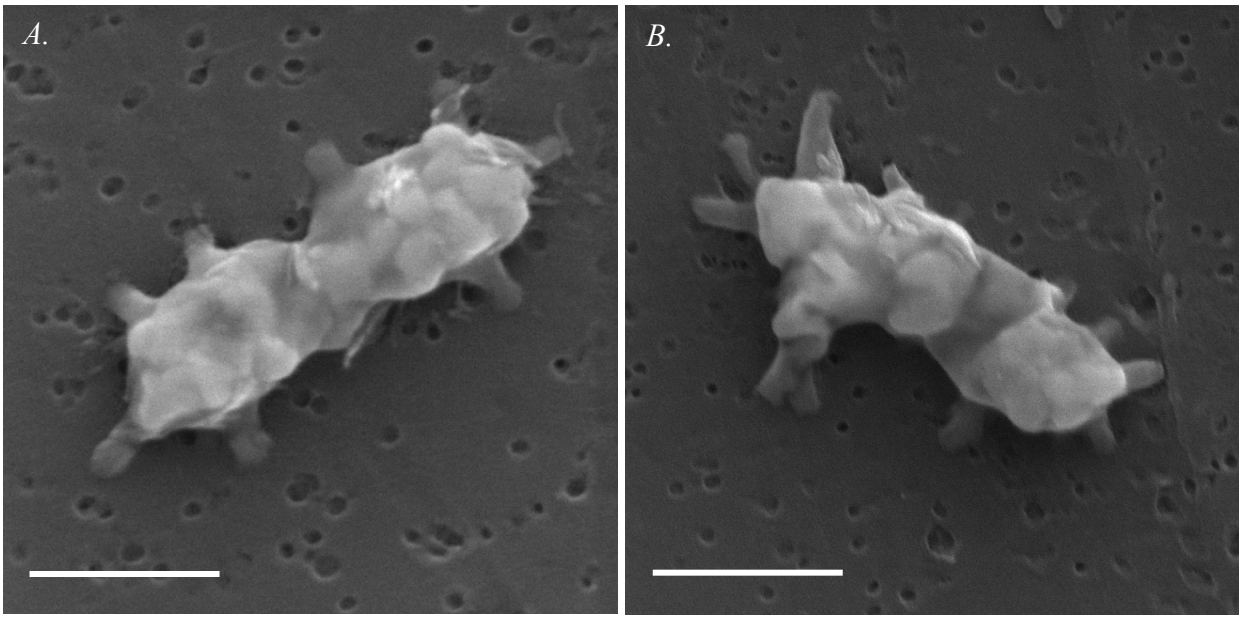




*Absorbance at 600 nm of mutant and wild type V. spinosum from a DAP-supplemented recovery. Samples included cultures transformed with (A) a cas9/guideRNA complex or no transformed sample or (B) the complex combined with a gene replacement or gene interruption cassette, cultured in M13 supplemented with kanamycin or M13 supplemented with DAP and kanamycin.*

### 7.3.2 SEM Imaging

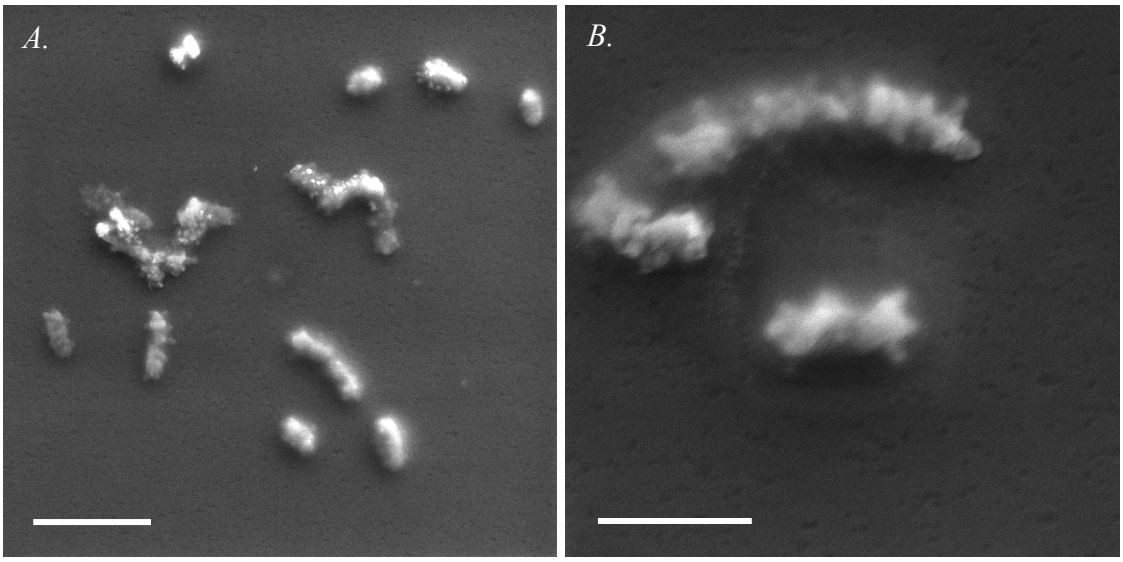
*V. spinosum* was successfully imaged between 60,000X and 80,000X magnification. Prosthecae were clearly visible on the surface of the wild type cells and the size ranged from 0.75 to 1.5 um in size (*Figure 15*). The regular morphology of most cells was an oblong double-ball shape, while some were only one ball (not shown). There was some irregularity in the imaging (*Figure 16*), however the majority of cells had consistent morphology in prosthecae and an oblong cell shape. Cells that congregated together still retained morphological definition and boundaries were easily discernible.

*Figure 15: Two hours post transformation of wild type cells*




*One cell each of wild type V. spinosum imaged at (A) 64,100X and (B) 51,600X magnification. Scale bars represent 1 um, each.*

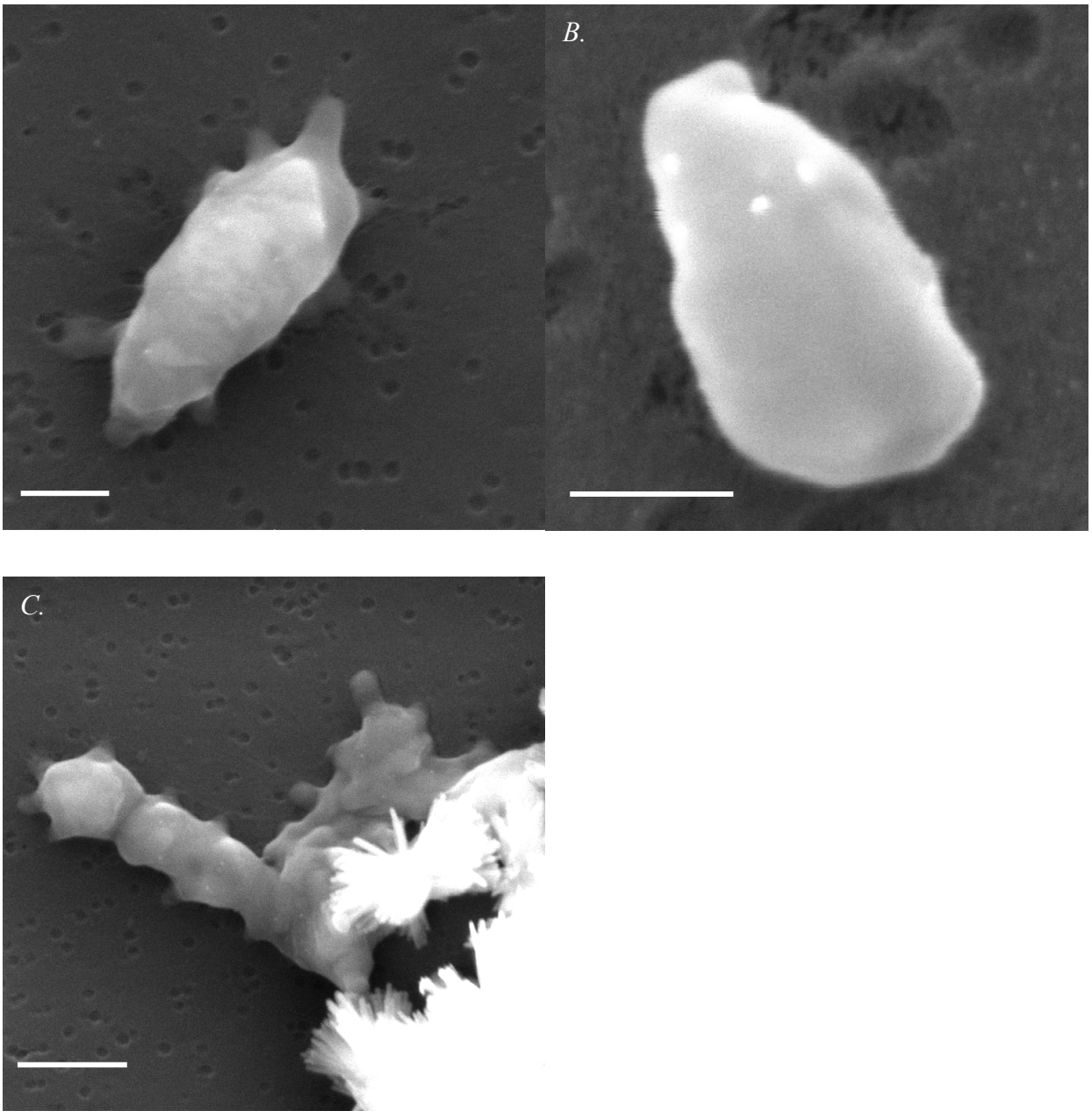
*Figure 16: Twenty-four hours post transformation of wild type cells*




*A cluster of (A) 20 wild type V. spinosum cells imaged at 8,720X magnification and (B) 6 wild type cells imaged at 27,900X magnification. Scale bars represent 2 um and 5 um, respectively.*

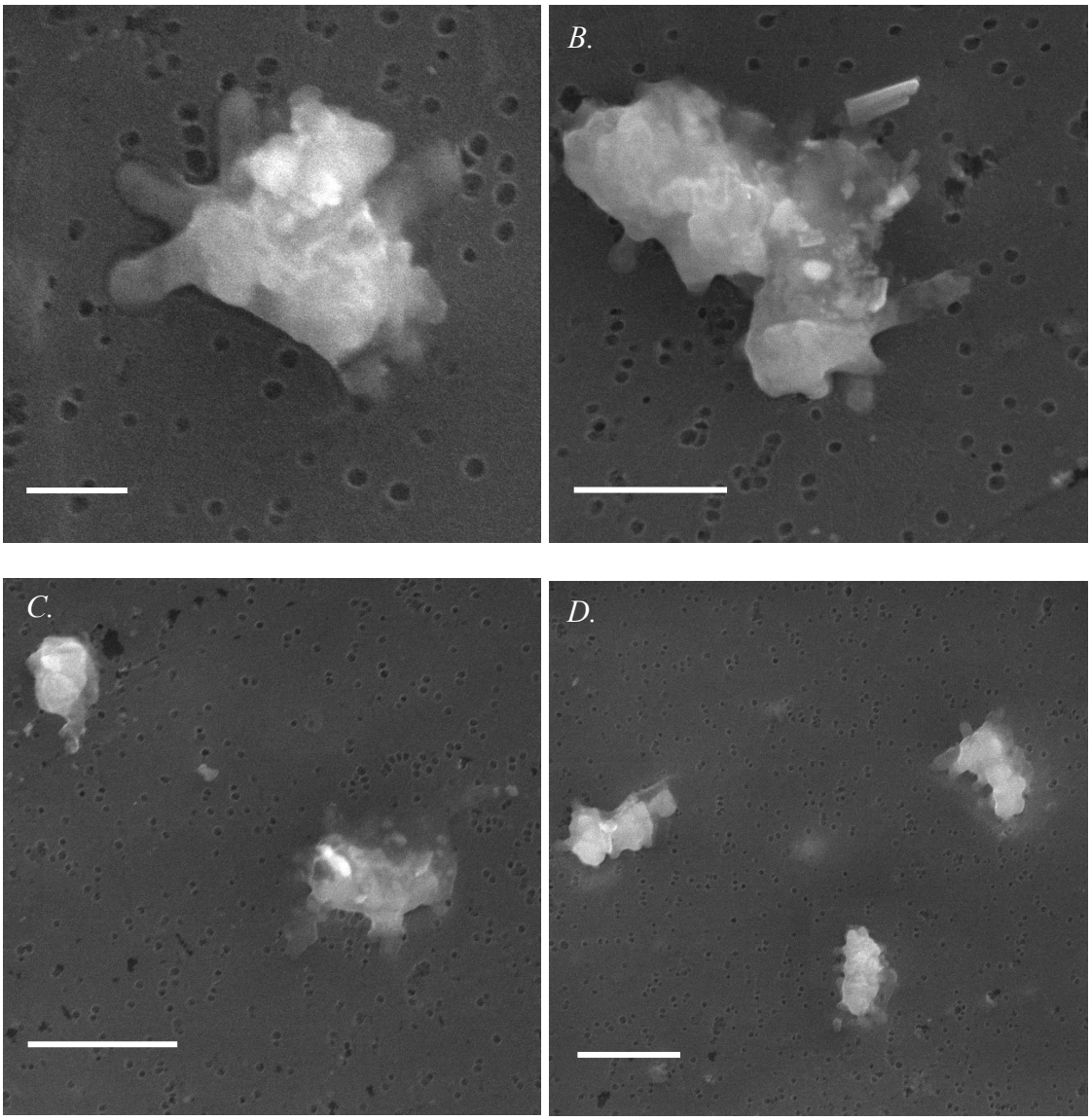
*V. spinosum* cells transformed with the cas9/gRNA targeting complex and gene interruption cassette were also successfully imaged after two and twenty-four hours of recovery. While the cellular sizes were consistently between 0.75 and 1.5 um, there was a wide range of morphological phenotypes present. After two hours (*Figure 17*), some cells were elongated and smooth with little to no protrusions or prosthecae on the surface. Other cells were amorphous, and indication of lysed cells was present. After twenty-four hours the phenotypic irregularity increased significantly (*Figure 18*). The surface composition of the cells exhibited more ridges and roughness and there were little to no typically oblong or double-ball morphologies present. Most cells were rounded and amorphous with signs of lysis and structural integrity failure.

*Figure 17: Two hours post transformation of mutant candidates from the cas9 complex and gene interruption cassette*




*(A) One V. spinosum cell at 67,500X magnification, scale bar at 500 nm. (B) V. spinosum cell at 311,000X magnification, scale bar at 200 nm. And (C) a cluster of V. spinosum cells at 41,900X magnification, scale bar at 1 um.*

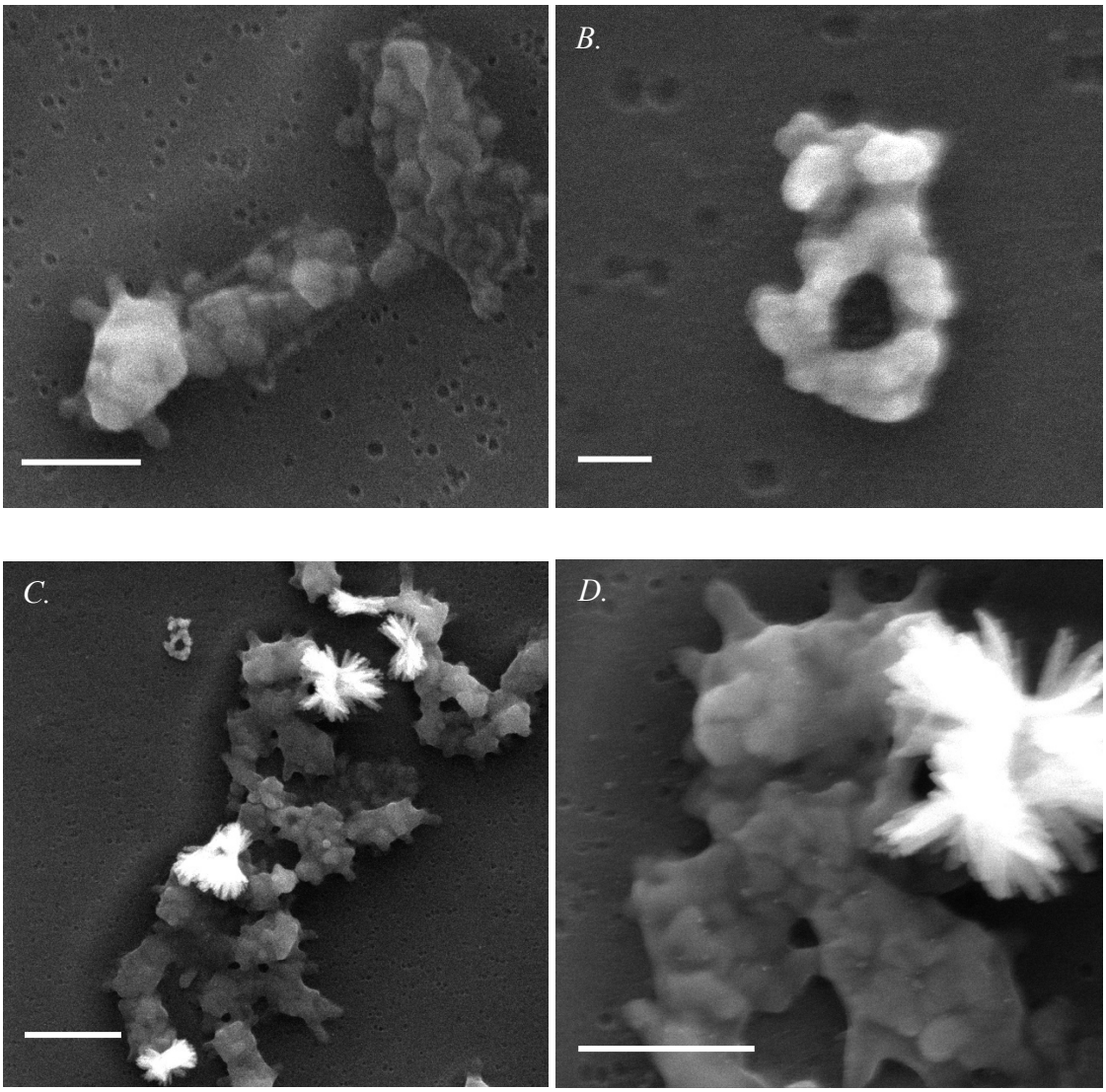
*Figure 18: Twenty-four hours post transformation of mutant candidates from the cas9 complex and gene interruption cassette*




*(A) One V. spinosum cell at 74,300X magnification, scale bar representing 500nm. (B) One V. spinosum cell at 57,400X magnification, scale bar respresenting 1 um. (C) Two V. spinosum cells at 27,800X magnification, scale bar representing 2 um, and (D) three V. spinosum cells at 19,800X magnification, scale bar representing 2 um.*

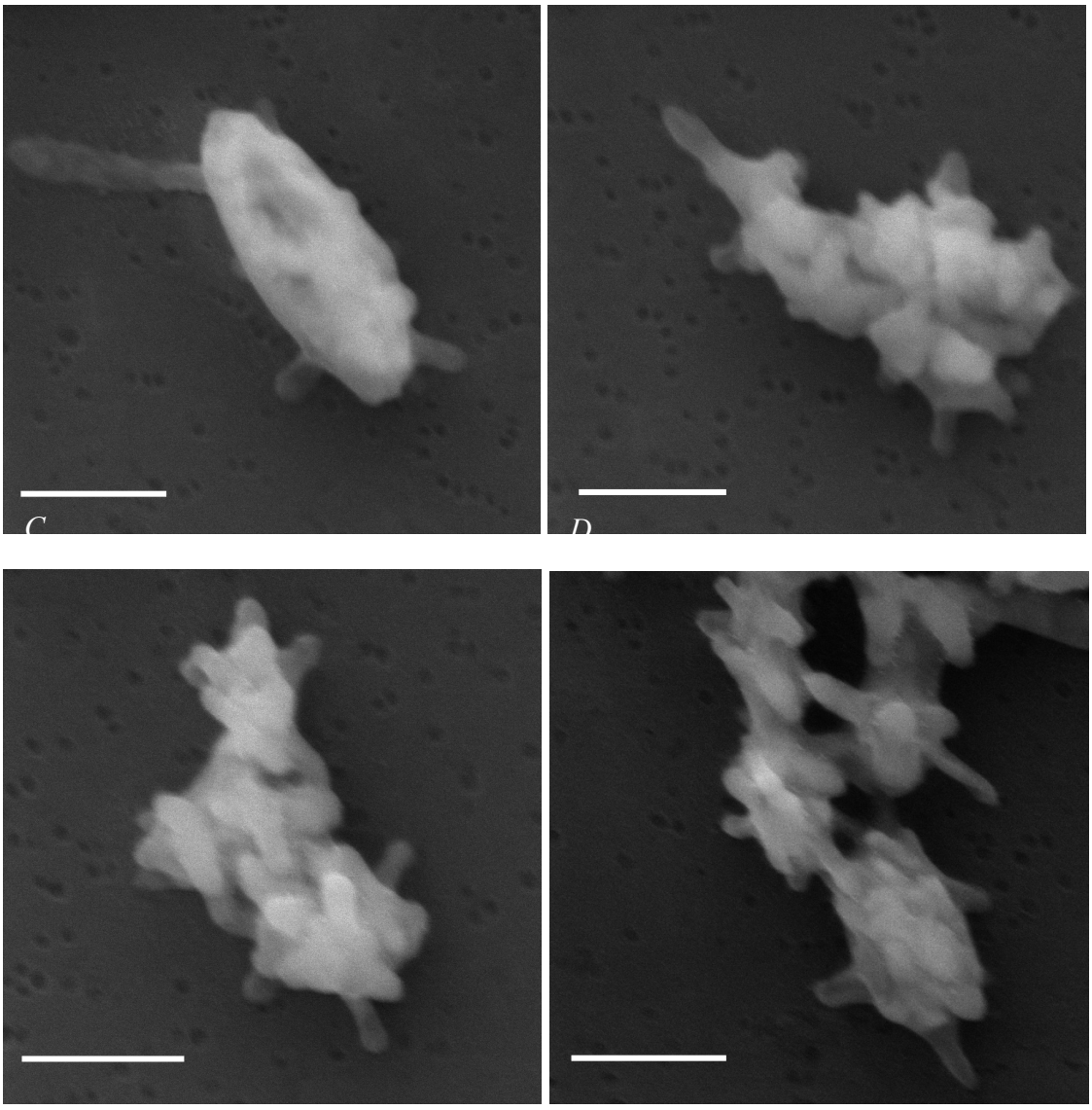
Cells transformed with the cas9/gRNA complex and gene replacement cassette were, once again, successfully imaged after two and twenty-four hours of recovery. Dramatic morphological irregularities were present. Two hours post-transformation, cells were seen largely congregated together (*Figure 19*). Most cells were amorphous and round, lacking the typical oblong shape and prosthecae. The surfaces of the cells were also rough and exhibited signs of folding and concavity. Remnants of cells were captured, indicating the possibility of lysis. Twenty-four hours post-transformation, evidence of phenotypic irregularity increased in frequency (*Figure 20*). Most of the imaged cells were fully unstructured, amorphous, and/or concaved. Cells that congregated together were so shapeless that definition and cellular boundaries were difficult to discern. Evidence of structural failure was also present.

*Figure 19: Two hours post transformation of mutant candidates from the cas9 complex and gene replacement cassette*




*(A) Two V. spinosum cells at 43,900X magnification, scale bar representing 1 um. (B) a cell fragment at 134,000X magnification, scale bar representing 200 nm. (C) a group of cells at 17,500X magnification, scale bar representing 2 um. (D) a group of cells at 65,200X magnification, scale bar representing 1 um.*

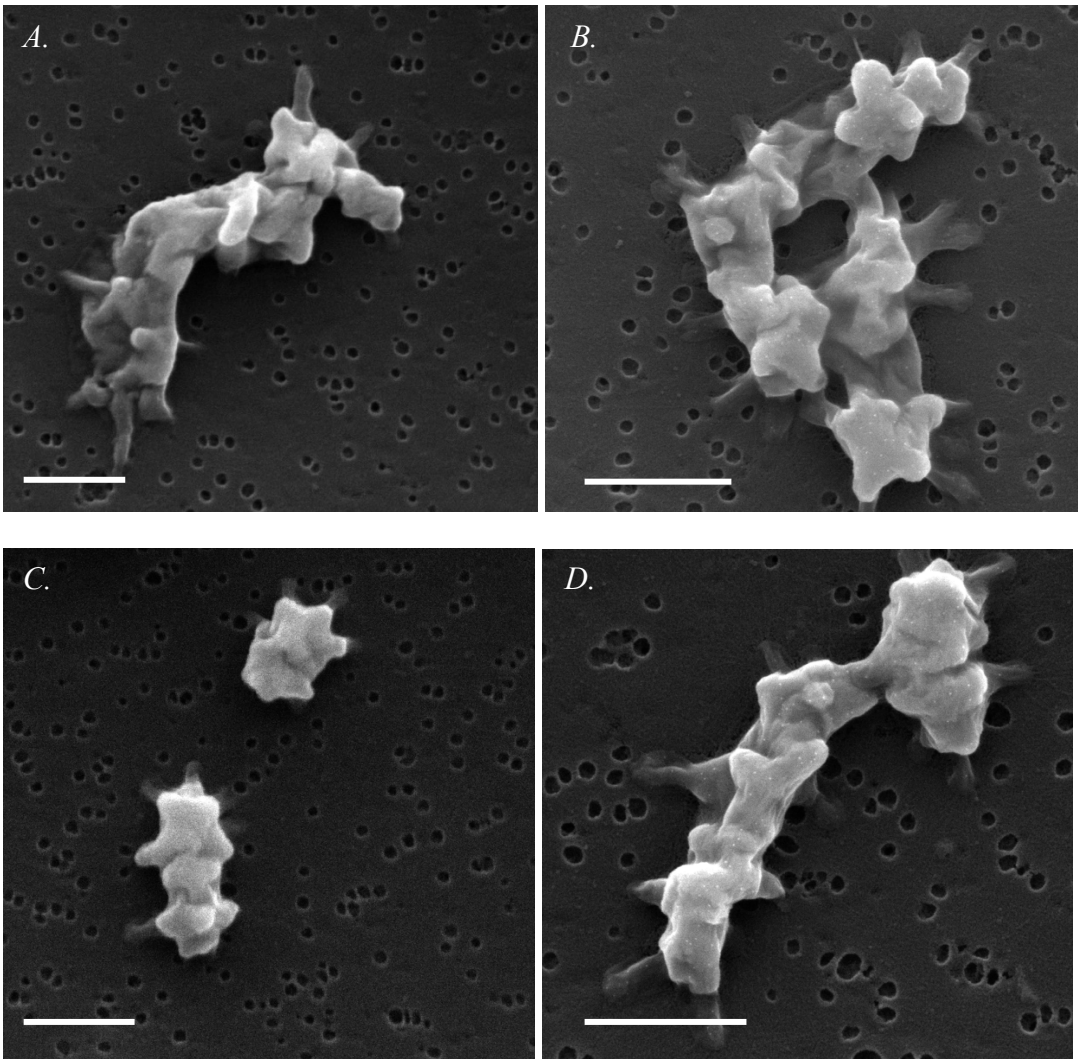
*Figure 20: Twenty-four hours post transformation of mutant candidates from the cas9 complex and gene replacement cassette*




*(A) One V. spinosum cell at 54,500X magnification, (B) a cell or cell group at 56,400X magnification, (C) a cell or cell group at 60,700X magnification, and (D) a group of cells at 57,200X magnification. Scale bars all represent 1 um.*

Imaging was only avaible twenty-four hours post-tranformation for cells transformed with excess amounts of the cas9/gRNA complex. Though the sample was limited to one time point, major phenotypic differences were once again present in high quality images (*Figure 21*). Cells formed long shapeless structures with no visible cellular definition aside from the exterior. The cell surface was concave and folded onto itself, indicating signs of structural failure. A handful of cells with normal morphology was found, but they were interpersed through the largely irregular and amorphous phenotypes.

*Figure 21: Twenty-four hours post transformation of mutant candidates from the cas9 complex*




*(A) One V. spinosum cell at 40,800X magnification, (B) a group of cells at 57,800X magnification, (C) two cells at 44,800X magnification, and (D) a cell or group of cells at 65,300X magnification. Scale bars each represent 1 um.*

## 7.4 DISCUSSION

*Verrucomicrobium spinosum* has unique growth characteristics requiring a low cell density in liquid culture and a humid environment for plating, characteristics that can hinder model system development. With the proper standardization, however, manipulation and development of this organism into a model is certainly possible. Growth to an $OD_{600}$ of ~0.1 for electrocompetent cell preparation, resuspension in equal volume post-electroporation, and plating of 100 ul results in the best plated phenotype with visible single colonies. The cells exhibited low transformation efficiency, as evidenced by the lack of successful transformations and lack of target genotypes or phenotypes, so alternative methods of transformation could be explored, including the use of conjugation instead of electroporation or increased cell concentration for electroporation.

When targeting *dapL* with a replication deficient vector and knockout cassette, the presence of colonies but lack of genotype, lack of phenotype, or lack of colonies altogether indicates that either no transformation, no recombination, or single recombination events were occurring. It can be theorized that a lack of secondary selection for the second recombination event combined with selection against recombinants via loss of the lysine biosynthetic pathway created a difficult environment for the integration of the gene. Any successfully recombined mutants were hypothetically instantly killed, or their growth inhibited and propagation of those mutants subsequently impossible. Only mutants with a functional *dapL* could grow, and those that also recombined the selectable marker could then propagate on antibiotic. Future work to generate a knockout via homologous recombination should include a secondary selectable marker, like a sucrose sensitivity gene, on the plasmid for selection against to then select for the second recombination event (Wu, Ren, & Huntley, 2015).

The *in vitro* CRISPR/cas9 system did show promise, however due to time limitations testing was halted at the culmination of the project. Transfection ready cas9 is optimized for eukaryotic cells, but these experiments show that it is also functional, and its use is realistic for bacteria that exhibit interesting properties or with unknown genetics. The presence of nuclear localization signals in the enzyme do not interfere with its activity upon the lack of a nucleus. In addition, the gRNA synthesis protocol was incredibly consistent and prolific, resulting in high yields of highly pure RNA for transformations. The consistency of the tool aided in developing a model system in an organism that is not widely studied and the tool has future potential to save time and money when targeting putatively essential genes.

In the case of DapL, though the naked controls successfully grew, a mutant was not recovered upon plating with for DAP and antibiotic selection. The organism's growth in liquid culture, however, showed signs of success. Though no mutants were isolated, and no growth was observed, the differential death rate of *V. spinosum* transformants in liquid culture was an indication of potential successful recombination of the selectable marker at the desired insertion site. In addition, scanning electron microscopy conducted in conjunction with knockout experiments indicates that *V. spinosum* was experiencing mutations that affected its phenotype. When transformed with the Cas9 complex, regardless of the presence or absence of a knockout cassette, cells became amorphous and congregated together in inseparable clumps. A qualitative investigation with SEM imaging also showed evidence of structural failure of the cell walls via lack of prosthecae, concavity, irregularity, and folds. The samples that included the knockout cassettes exhibited a higher frequency of highly irregular cellular phenotypes, with more frequent and severe structural failures. The presence of atypical cells in the test cases indicates potential failures in cell wall integrity, perhaps resulting from the failure of peptidoglycan

crosslinking. Based on this evidence, recovery of a mutant may not be possible. However, when combined with the differential growth of transformants in liquid culture, the data do support the essentiality of *dapL* and the efficacy of it as an antibiotic target.

Future work in evaluating VsDapL (as a model for any DapL enzyme) as an antibiotic target should account for the gaps detailed above. First, a vector that definitively replicates in *V. spinosum* should be identified, as complementing the organism with a functional *dapL* prior to a knockout would mitigate the extremely toxic effects indicated by previous experiments. A method to explicitly test the efficacy of the Cas9 system should be developed; the *dapL* sequence on a vector in *E. coli* could be targeted with the same gRNA to accommodate any transformation-associated issues and directly test the Cas9 system. In addition, dynamic approaches to measure a knockdown or knockout in liquid culture, like RNA-interference, should be explored to accommodate the growth properties of *V. spinosum*. It would also provide a refined view of the toxicity of inhibiting DapL without eliminating the functionality of the gene and live/dead staining and fluorescence microscopy could be utilized to provide further insight into the effect of the knockout on individual cells. Lastly, future work should assess transformation efficiency and recovery methodology in *V. spinosum* to clarify these conclusions and provide more insight.

Overall, the experiments to generate a genetic knockout of *dapL* in *V. spinosum* with the goal of using it as a model in the development of narrow spectrum antibiotics were largely exploratory. They consisted of extensive troubleshooting, trial, and error. However, they did succeed in identifying tools and developing standardized protocols in the growth, preparation, and manipulation of *V. spinosum* for use in future experiments and laid the foundation for supporting the essentiality of *dapL* in the organism and, by association, all bacteria that employ the pathway.

# 8 CHAPTER 2: Comparative Molecular Dynamics Simulations of DapL

## 8.1 INTRODUCTION

Computational biology is a rapidly expanding field that allows researchers to experiment *in silico*; they can perform simulations and construct models of biological molecules and processes to gain insight prior to, during, and after experimentation. Molecular dynamics, a field of computational chemistry, is rapidly expanding with application in computational biology. In short, it applies molecular mechanics methods to predict the motion of a molecule in an environment over a time course. Molecular mechanics, approaching molecules as collections of balls and springs to represent the atoms and bonds respectively, calculates the potential energy of the molecular system and readjusts the position of each atom until it obtains the lowest energy state (Lewars, 2003).

Molecular dynamics applies Newton's laws of motion and molecular mechanics in a stepwise manner, calculating the different forces exerted on each atom and adjusting the position for energy minimization over time, usually on a femtosecond time scale. The forces exerted on an atom can be defined in bond stretching, angle bending, torsion, van der Waals potentials, or electrostatic potentials and are generally referred to as defined force fields of inter- and intra-molecular interactions for a given molecule type or category (Jensen, 2007). There are numerous molecular dynamics packages available for use, one of such is titled "AMBER" (Assisted Model Building with Energy Refinement). It is both a molecular dynamics package and a set of pre-determined molecular mechanical force fields (tested and determined parameters that dictate the interaction of biomolecules) that can be used to simulate biomolecule movement (CASE et al., 2005). At the culmination of the molecular dynamics calculations, the stepwise updated atomic coordinates, velocities of movement, and accelerations are processed to generate trajectories that,

when combined with statistical mechanics, describe the time evolution of the molecular system. In computational biology, molecular dynamics can be applied to nearly anything including proteins, DNA, and RNA (with limitations of size and stability) and, when combined with other techniques like protein docking, can provide novel insight into biomolecule form and function.

Protein docking is a molecular modeling technique that non-covalently associates a molecule (usually a ligand) and a protein to determine a three-dimensional complex. The interaction between two molecules is complex; the technique must account for the complexities of the protein (e.g. amino acid interactions and hydrophobicity), the molecule (e.g. flexibility and charge), and the combination of the two (e.g. electrostatic forces and hydrogen bonds). In order to model this interaction, docking protocols utilize a combination of a search algorithm and scoring function. The search algorithm creates a set of molecular configurations that is evaluated by the scoring function, a number of mathematical models that predict the strength of each interaction in the set. Depending on the size of the system and computational limitations, different search algorithms and scoring functions can be used. For larger molecules, docking is limited to molecular mechanics but others may employ Monte Carlo methods or genetic algorithms combined with various scoring functions including knowledge-based or consensus scoring (Hernández-Santoyo, Tenorio-Barajas, Altuzar, Vivanco-Cid, & Mendoza-Barrera, 2013; Jensen, 2007). One of such docking programs, AutoDock Vina has a largely computationally-optimized scoring function and a search algorithm that requires a user-defined search space for fast and accurate processing (Trott & Olson, 2010). Though complex, docking has various applications, including structure-guided drug design and *in vitro* protein-protein or protein-nucleic acid interactions.

Here, a comprehensive, comparative molecular dynamics simulation package, DROIDS (Detecting Relative Outlier Impacts in Dynamic Simulations) 2.0 (Babbitt, Mortensen, Coppola, Adams, & Liao, 2018), was used in conjunction with AutoDock Vina to provide insight into the binding dynamics of putative inhibitory lead compounds. Developed as a GPU-based package, DROIDS can be used on any linux computer containing graphics cards (e.g. gaming computers or specifically designed research computers). Because of advances in graphics card development, molecular dynamics calculations can now be performed at speeds orders of magnitude faster and biologically relevant calculations are now realistic expectations. Utilizing Amber16 and Ambertools17 packages (CASE et al., 2005), cpptraj vector trajectory analysis (Roe & Cheatham, 2013), the R software suite for statistical analysis, and UCSF Chimera visualization tools (Pettersen et al., 2004) in GUI-based pipeline, DROIDS calculates the local modes of interaction between the residues of a given structure, simulating the interaction between each molecule, for two macromolecules over a set time course for a user-designated number of replications. It then compares the dynamics of the two macromolecules and presents a statistical analysis of the differences in dynamic movement in the context of atomic fluctuations from the mean position. Because this information is presented in graphs, figures, and structure coloration, there is great potential for interpretation and dissemination across a broad audience. In addition, the implementation of DROIDS in a user-interface platform with accessible graphics-card based system dependencies harnesses usability for any given user. The methods and results presented here can be applied for the basic scientist and modified to not only supplement laboratory experiments, but also provide a rational guide in their development.

Because of its presence in key medically and biotechnologically relevant organisms, DapL is an attractive target for antibiotic, herbicide, and algacide development. Previous work identified

antibiotic lead compounds in a comprehensive screening analysis that inhibited the DapL enzyme from *A. thaliana, L. interrogans, C. reinhardtii,* and *V. spinosum.* Four compounds specifically effectively inhibited the activity of the *V. spinosum* DapL (VsDapL) enzyme, derived from four classes with different structural elements (McKinnie et al., 2014). However, the analysis relied solely on experimental data and manual screening of a 30,000-compound library and an alternative method to screen the database to a smaller set of compounds would be beneficial.

To date, the crystal structure of VsDapL has not been published. However, due to sufficient sequence similarity with related DapL enzymes (i.e. *Arabidopsis thaliana* PDB: 3EI7*, Chlamydia trachomatis* PDB: 3QGU*,* and *Chlamydomonas reinhardtii* PDB: 3ASA) it is possible to construct a homology-based model and identify active-site residues. When compared to the unpublished, experimentally determined structure from Anthony Weatherhead at the University of Canterbury, the model is sufficient for analyses (RMSD of 1.15 and Q-Score of 0.825) to make inferences about the form, function, and activity of the enzyme. To demonstrate the scope and applicability of the DROIDS pipeline, test the accuracy of DROIDS in drawing conclusions, and provide refined insight into the screening analysis (McKinnie et al., 2014), and demonstrate the use of a modeled enzyme in such analyses, the effect of the four identified inhibitory lead compounds on the molecular dynamics of the VsDapL model was analyzed and trends in the resulting dynamics change were identified.

## 8.2 METHODS

### 8.2.1 Multiple Sequence Alignment

Multiple sequence alignment was constructed in MEGA of the DapL protein sequence from *V. spinosum* (PDB: WP_09961032.1)*, A. thaliana* (UniProt: Q93ZN9)*, C. trachomatis* (UniProt: G4NMX8)*,* and *C. reinhardtii* (UniProt: A8IW39). Sequences were aligned via MUSCLE algorithm. Conserved active site loops and residues were identified from the multiple sequence alignment.

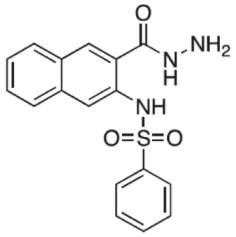
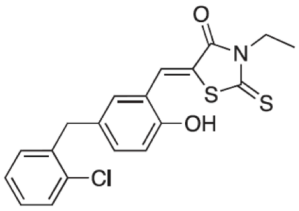
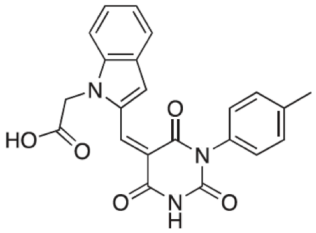
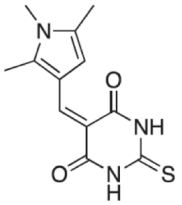
### 8.2.2 Structure Stabilization Testing

Structure stability and conformation testing was performed by Jamie Mortensen at RIT during the initial DROIDS pipeline development and prior to version 1 release. The solvated structures were each subjected to 1000 cycles of minimization using a steepest-descent algorithm, followed by an 80-picosecond heating step where the temperature was gradually raised from 0 to 300K and then an 80-picosecond simulation allowing for the potential energy of the molecule to reach an equilibrium, followed by a 120-nanosecond sampling simulation. Atomic coordinates were calculated every 0.002 picoseconds and recorded every 0.2 picoseconds. Temperature and pressure were held constant for the equilibration and subsequent runs.

### 8.2.3 Compound Preparation

Lead compounds identified by (McKinnie et al., 2014) (*Figure 22*) were exported in SMILES format from the Chembridge database ("ChemBridge | Home," n.d.) and the 3D coordinates were generated with the CORINA server (Sadowski, Gasteiger, & Klebe, 1994; Schwab, 2010; "Welcome to MN-AM | MN-AM," n.d.).

*Figure 22: Antagonist Lead Compounds Identified for Use*

| Compound | ChemBridge ID Number | Chemical Formula | Structure |
|---|---|---|---|
| Hydrazide | 5925714 | $C_{17}H_{15}N_3O_3S$ |  |
| Rhodanine | 6523070 | $C_{19}H_{16}ClNO_2S_2$ |  |
| Barbiturate | 6072466 | $C_{22}H_{17}N_3O_5$ |  |
| Thiobarbiturate | 6088649 | $C_{12}H_{13}N_3O_2S$ |  |

*The ID number, chemical formula, and structure for the four antagonistic lead compounds against DapL identified by and adapted from McKinnie et al. and assigned for use in the molecular dynamics analyses.*

### 8.2.4 Molecular Docking

Three dimensional compounds were docked onto the DapL active site with the PyRX interface for Autodock Vina (Dallakyan & Olson, 2015; Trott & Olson, 2010). The best ranked conformation, with the best binding affinity reported, was chosen for each respective molecular dynamics calculation run in DROIDS. For each run, the protein-ligand complex, the protein, and the ligand in its proper conformation identified by Vina were exported in PDB format. The resulting three files each were used.

### 8.2.5 DROIDS Settings and Run Times

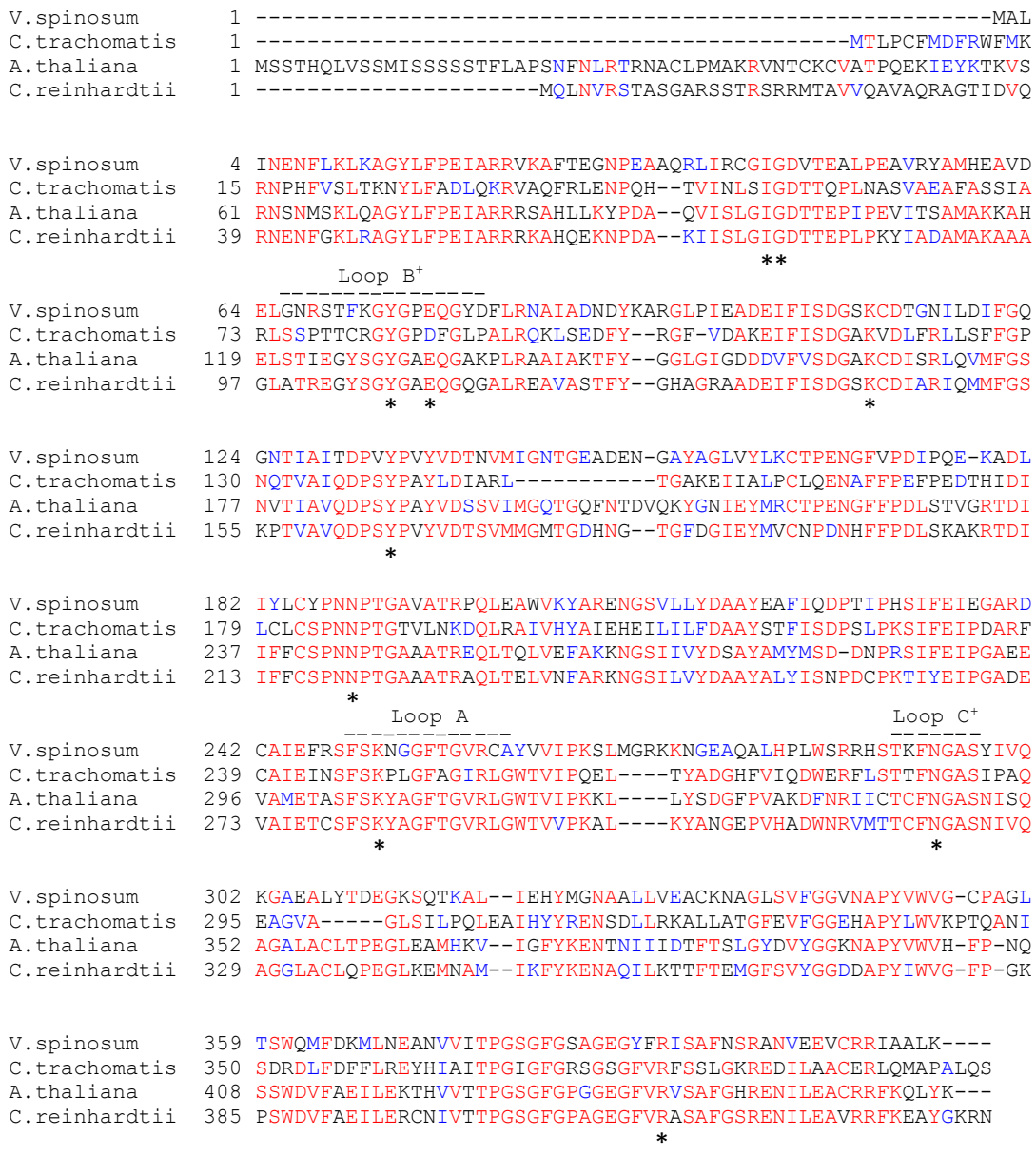
The structures solvated structures were each subject to 1000 cycles of minimization using a steepest-descent algorithm, followed by a 100-picosecond heating step to 300K, followed by a 5-nanosecond equilibration step to allow for the potential energy of the molecule to reach an equilibrium before 300 production sampling runs of 0.5 nanoseconds. Temperature and pressure were held constant for the equilibration and subsequent runs.

## 8.3 RESULTS

### 8.3.1 Multiple Sequence Alignment

Loops lining the active site in *V. spinosum* were identified between residue numbers 249 and 261 (Loop A), as well as those from the opposing chain between residues 66 and 81 (Loop B), and 291 and 297 (Loop C). In addition, key conserved active site residues were identified as residue numbers 43, 44, 74, 77, 111, 134, 189, 251, 294, 390 (*Figure 23*).

*Figure 23: VsDapL Multiple Sequence Alignment*

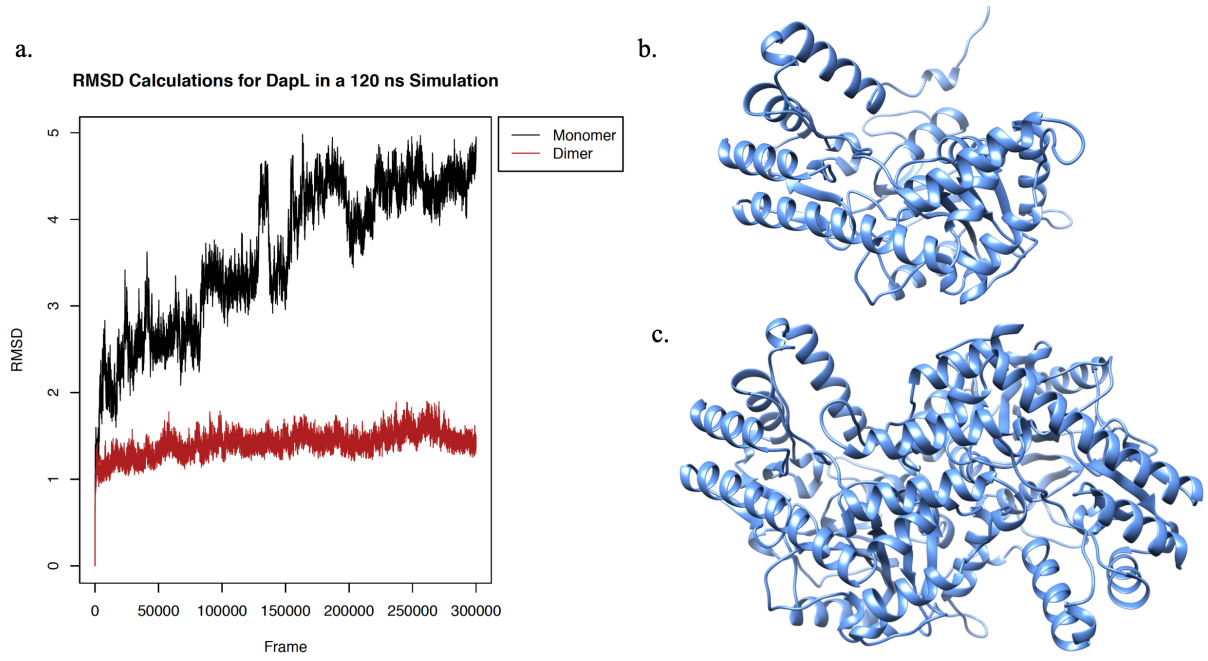



*Alignment of the VsDapL sequence to reference sequences. * denotes conserved active site residues, "-----" denotes residues in loops lining the active site, and + denotes a contribution from the opposing chain.*

### 8.3.2 Structure Stability Testing

Based on RMSD (Root Mean Square Deviation) calculations of the fluctuation of the alpha-carbons of each amino acid from their mean positions, the homodimer and monomer forms of the DapL protein were compared (*Figure 24*). As a monomer, the early sampling of vsDapL coordinates indicate a RMSD value of 2. Over time, however, the RMSD increases and nears 5. On the contrary, the vsDapL homodimer stabilizes just above a RMSD value of 1 for the duration of the sampling time.

*Figure 24: VsDapL Dimer and Monomer Testing*




*A. Root Mean Square Deviation (RMSD) values for the fluctuation of the alpha-carbons of each amino acid from the mean position of the backbone were calculated for the molecular dynamics sampling runs of the B. DapL monomer and C. DapL homodimer.*

### 8.3.3 Molecular Docking

Upon docking, the hydrazide molecule settled in a best conformation with a binding affinity of -7.6 kcal/mol. There was a total of five conformations found, all associating in various positions across the entire protein as indicated by the RMSD values when compared to the best identified conformation. The binding affinity ranged from -7.6 kcal/mol to -5.2 kcal/mol. The molecular structure for the best docked conformation indicated the hydrazide molecule bound close to the active site (*Figure 25; Table 1*).

The rhodanine molecule settled in a best conformation with a binding affinity of -8.3 kcal/mol. There was a total of nine conformations found, all associating in various positions across the entire protein as indicated by the RMSD values when compared to the best identified conformation. The binding affinity ranged from -8.3 kcal/mol to -6.7 kcal/mol. The molecular structure for the best docked conformation indicated the rhodanine molecule docked in the minor arm of one active site (*Figure 26; Table 2*).

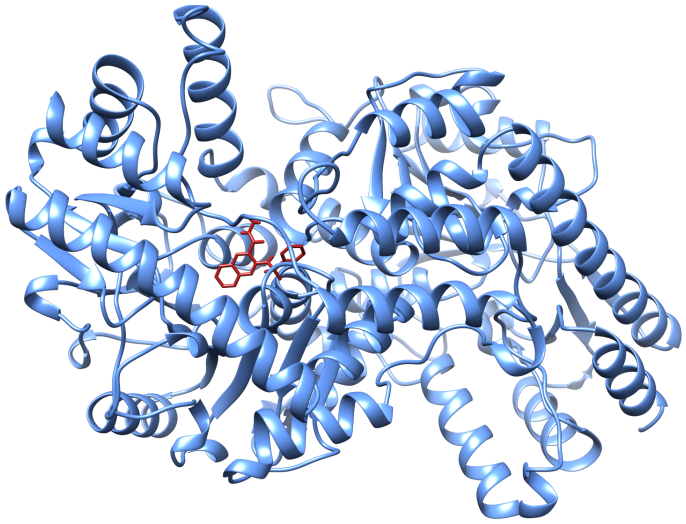
The barbiturate molecule settled in a best conformation with a binding affinity of -7.6 kcal/mol. There was a total of nine conformations found, all associating in various positions across the entire protein as indicated by the RMSD values when compared to the best identified conformation. The binding affinity ranged from -7.6 kcal/mol to -7.2 kcal/mol. The molecular structure for the best docked conformation indicated the barbiturate molecule docked along the outer edge near the active site (*Figure 27; Table 3*).

The thiobarbiturate molecule settled in a best conformation with a binding affinity of -6.4 kcal/mol. There was a total of nine conformations found, all associating in various positions across the entire protein as indicated by the RMSD values when compared to the best identified conformation. The binding affinity ranged from -6.4 kcal/mol to -5.7 kcal/mol. The molecular

structure for the best docked conformation indicated the thiobarbiturate molecule docked on the outer surface of the protein (*Figure 28; Table 4*).

Based on identified contacts, the four small molecule antagonists bound to different locations on DapL. Hydrazide bound to the majority of residues in the active site, including the catalytic lysine (residue 251). Rhodanine and thiobarbiturate did not bind near the active site whatsoever, and barbiturate bound near the active site but not directly in it. Rhodanine docking resulted in the best binding affinity of the four compounds. Docking of hydrazide and barbiturate resulted in similar binding affinities, and thobarbiturate resulted in the least impactful docking (*Figure 29)*.

*Figure 25: Hydrazide Docking Structure*



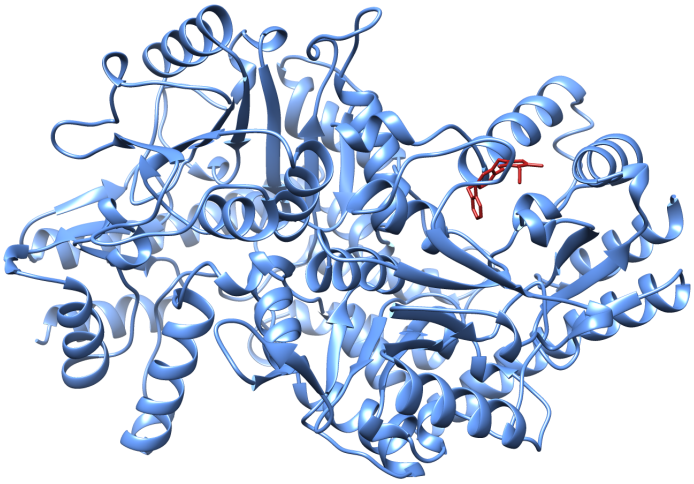
*The structure of one hydrazide molecule (red) docked near one active site of VsDapL (blue).*

*Table 1: Hydrazide Docking Statistics*

| Ligand | Binding Affinity (kcal/mole) | rmsd/ub | rmsd/lb |
|---|---|---|---|
| VsDapL_1 | -7.6 | 0 | 0 |
| VsDapL_2 | -7.4 | 7.064 | 2.225 |
| VsDapL_3 | -5.5 | 6.538 | 2.595 |
| VsDapL_4 | -5.3 | 26.033 | 23.535 |
| VsDapL_5 | -5.2 | 2.206 | 1.78 |

*Docking results for the hydrazide and VsDapL association. The lowest energy structure is reported first, and the RMSD of the following structures is reported as compared to the position of the first conformation.*

*Figure 26: Rhodanine*



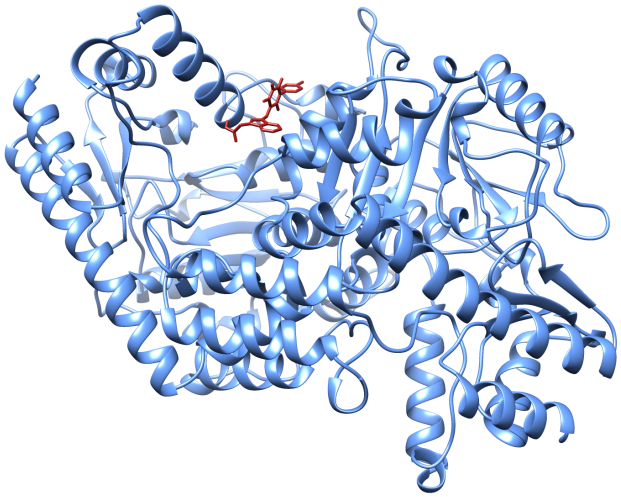
*The structure of one rhodanine molecule (red) docked near one major arm of VsDapL (blue).*

*Table 2: Rhodanine Docking Statistics*

| Ligand | Binding Affinity (kcal/mole) | rmsd/ub | rmsd/lb |
|---|---|---|---|
| VsDapL_1 | -8.3 | 0 | 0 |
| VsDapL_2 | -8.1 | 2.053 | 1.579 |
| VsDapL_3 | -7.6 | 3.294 | 2.564 |
| VsDapL_4 | -7.2 | 5.09 | 3.438 |
| VsDapL_5 | -7.1 | 2.24 | 1.756 |
| VsDapL_6 | -6.9 | 18.955 | 16.319 |
| VsDapL_7 | -6.8 | 4.915 | 3.176 |
| VsDapL_8 | -6.7 | 16.581 | 15.273 |
| VsDapL_9 | -6.7 | 4.339 | 2.866 |

*Docking results for the rhodanine and VsDapL association. The lowest energy structure is reported first, and the RMSD of the following structures is reported as compared to the position of the first conformation.*

*Figure 27: Barbiturate*



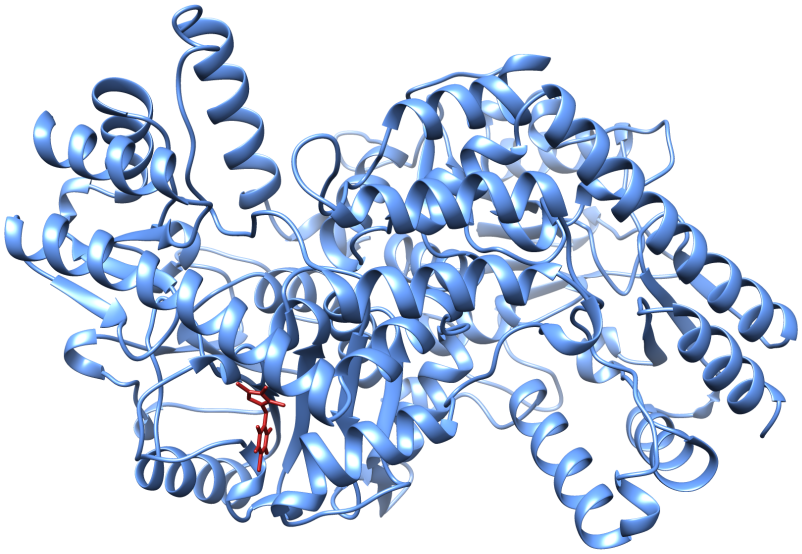
*The structure of one barbiturate molecule (red) docked near one active site of VsDapL (blue).*

*Table 3: Barbiturate Docking Statistics*

| Ligand | Binding Affinity (kcal/mole) | rmsd/ub | rmsd/lb |
|---|---|---|---|
| VsDapL_1 | -7.6 | 0 | 0 |
| VsDapL_2 | -7.5 | 9.86 | 7.352 |
| VsDapL_3 | -7.4 | 8.358 | 2.628 |
| VsDapL_4 | -7.4 | 3.049 | 1.801 |
| VsDapL_5 | -7.3 | 9.42 | 7.243 |
| VsDapL_6 | -7.3 | 5.406 | 3.305 |
| VsDapL_7 | -7.3 | 5.49 | 3.092 |
| VsDapL_8 | -7.2 | 23.098 | 17.59 |
| VsDapL_9 | -7.2 | 5.499 | 2.922 |

*Docking results for the barbiturate and VsDapL association. The lowest energy structure is reported first, and the RMSD of the following structures is reported as compared to the position of the first conformation.*

*Figure 28: Thiobarbiturate*



*The structure of one thiobarbiturate molecule (red) docked near the bottom of one VsDapL active site (blue).*

*Table 4: Thiobarbiturate Docking Statistics*

| Ligand | Binding Affinity (kcal/mole) | rmsd/ub | rmsd/lb |
|---|---|---|---|
| VsDapL_1 | -6.4 | 0 | 0 |
| VsDapL_2 | -6.3 | 2.188 | 1.166 |
| VsDapL_3 | -6.2 | 4.465 | 3.288 |
| VsDapL_4 | -6.2 | 30.347 | 29.039 |
| VsDapL_5 | -6.1 | 30.138 | 28.76 |
| VsDapL_6 | -5.9 | 26.064 | 24.954 |

| | | | |
|---|---|---|---|
| VsDapL_7 | -5.9 | 26.127 | 25.066 |
| VsDapL_8 | -5.8 | 27.866 | 25.529 |
| VsDapL_9 | -5.7 | 29.549 | 28.307 |

*Docking results for the thiobarbiturate and VsDapL association. The lowest energy structure is reported first, and the RMSD of the following structures is reported as compared to the position of the first conformation.*

*Figure 29: Binding Summary*

| Compound | Binding Affinity (KD) | Interacting Residues |
|---|---|---|
| Hydrazide | -7.6 | 43*, 74*, 111*, 110, 134*, 189*, 217, 250 , 251**, 259, 294* |
| Rhodanine | -8.3 | 12, 15, 16, 135, 138, 156, 161, 361, 379 |
| Barbiturate | -7.6 | 17, 18, 21, 77*, 78, 287, 290 |
| Thiobarbiturate | -6.4 | 91, 95, 221, 231, 247, 316, 317 |

* Active site residue
** Catalytic residue

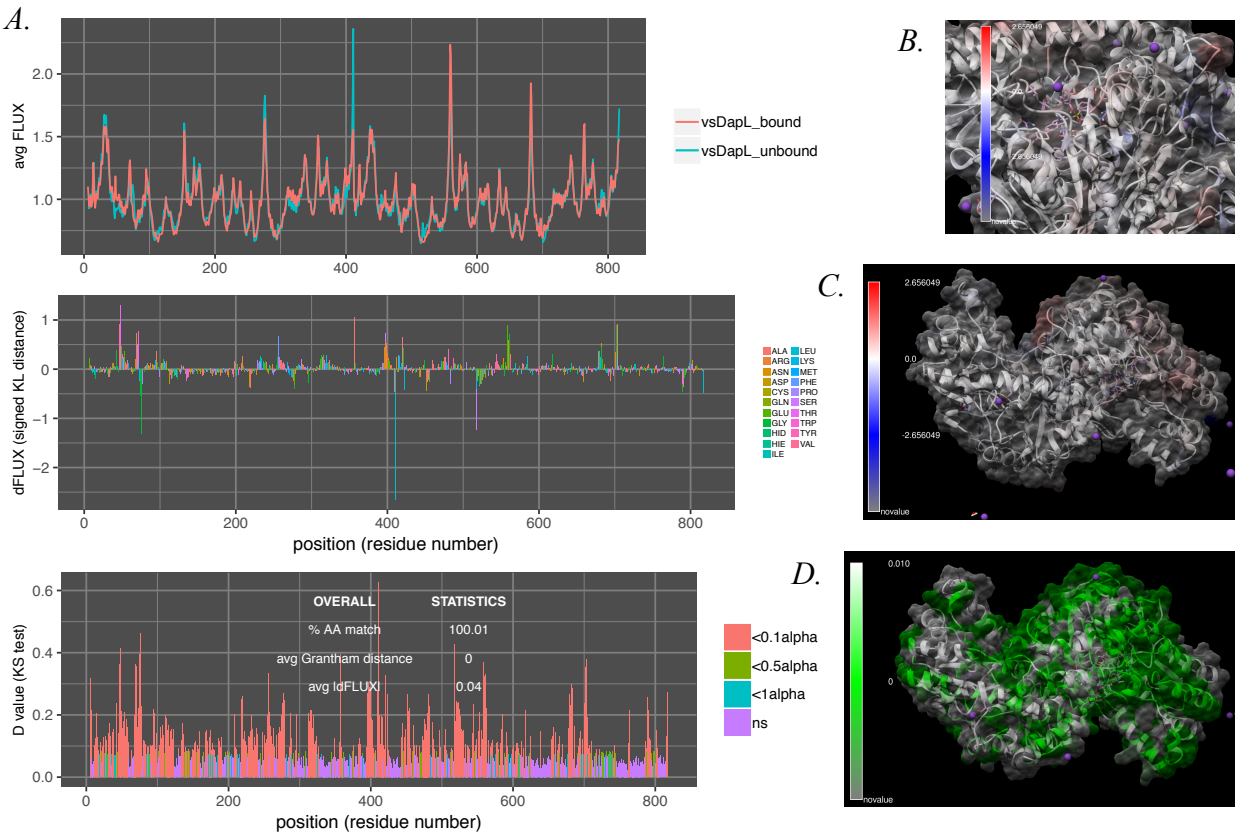
*The binding affinity for and residues that interact with each identified compound used in this study. *Residues that are associated with the active site. **Catalytic residue.*

### 8.3.4 Comparative Molecular Dynamics

Statistically significant differences in fluctuations between the unbound VsDapL molecule and VsDapL bound to hydrazide were found upon completion of the DROIDS pipeline (*Figure 30*). The fluctuation (dFLUX) and statistical profiles were asymmetric across the protein, they were not replicated between the two subunits (e.g. the dFLUX of residues 1-10 on chain A was different than the dFLUX of residues 1-10 on chain B). A full list of residues that experienced statistically significant differences in fluctuation are presented in *Figure 38.* The catalytic lysine

in the active site corresponding to the bound molecule was dampened and its activity stabilized, and the catalytic loop in both pockets was destabilized as some conserved residues were dampened while others were softened with an amplification in fluctuation. In fact, every active site chain was destabilized upon hydrazide association, some residues in each active site loop were dampened while others were activated. The statistically significant differences in fluctuation were primarily along the active site and dimer interface.

*Figure 30: DROIDS Plot and Images, VsDapL and Hydrazide*

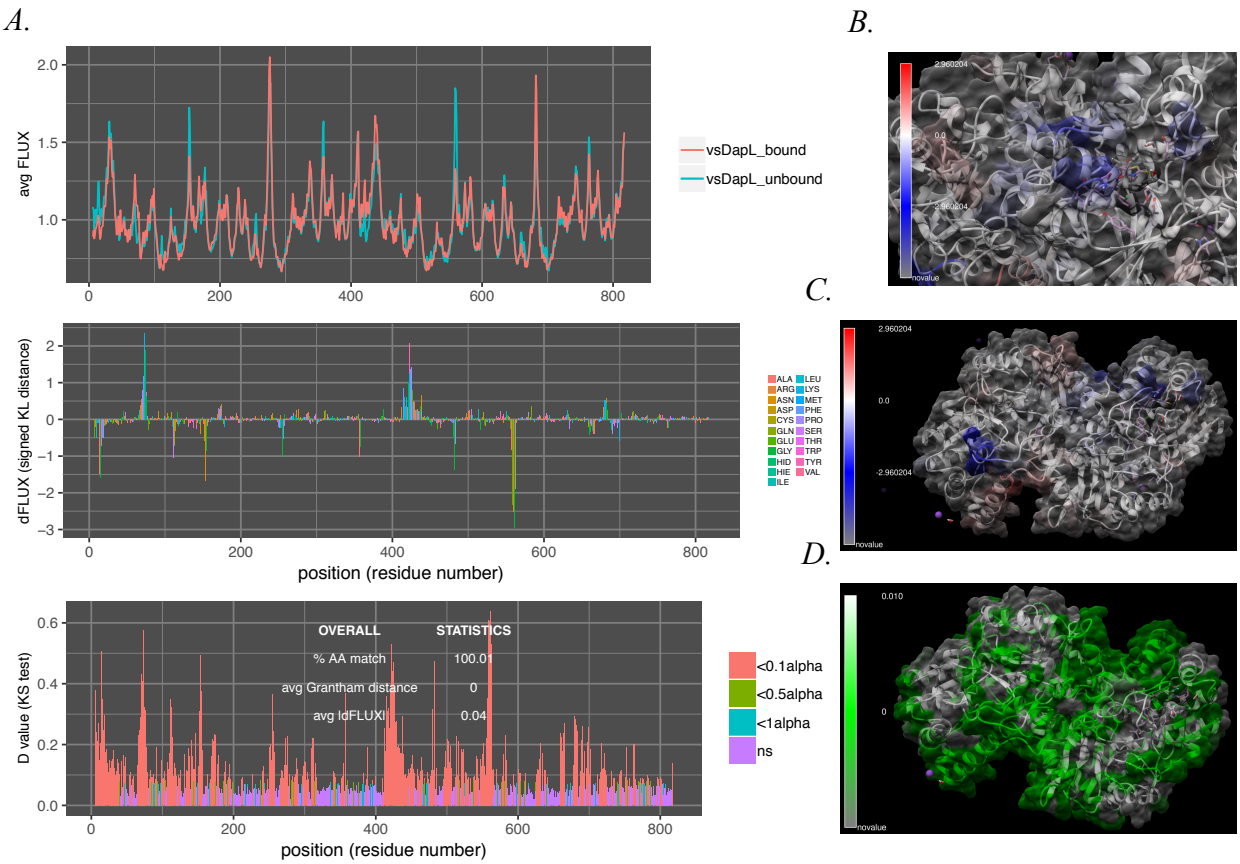



*300 x 0.2ns explicit solvent molecular dynamic simulation of wild type V. spinosum DapL enzyme compared to DapL associated with a hydrazide molecule. (A) Average atom fluctuation profiles (top plot), signed symmetric Kullback-Leibler (KL) divergences in local atom fluctuation distributions, collected from the N, CA, C and O atoms from each amino acid on the polypeptide*

*backbone (middle plot), and P values from a Benjamini-Hochberg corrected Kolmogorov-Smirnov (KS) test indicating significant differences in dynamics is also shown (bottom plot). The KL divergences in local atom fluctuation distributions are shown color mapped to temperature on the protein structure with red indicating softer regions with amplified fluctuation and blue indicating stabilized regions with dampened fluctuation (B) across the entire structure and (C) on the active site. (C) P values from the KS test indicating significant differences in dynamics.*

Statistically significant differences in fluctuations between the unbound VsDapL molecule and VsDapL bound to rhodanine were also found upon completion of the DROIDS pipeline (*Figure 31*). The fluctuation (dFLUX) and statistical profiles were asymmetric across the protein, they were not replicated between the two subunits (e.g. the dFLUX of residues 1-10 on chain A was different than the dFLUX of residues 1-10 on chain B). A full list of residues that experienced statistically significant differences in fluctuation are presented in *Figure 38.* The catalytic lysine in the active site corresponding to the bound molecule was dampened and its activity stabilized, and the conserved catalytic loop in that pocket was destabilized as some residues were dampened while others were softened with an amplification in fluctuation. The catalytic loop in the second pocket (not corresponding to the location of the bound compound) was significantly dampened. At least one residue on each active site loop was dampened, but some loops did not have any corresponding amplification in activity. The statistically significant differences in fluctuation were broadly spread across the protein, not limited to any region or interface. However, both active sites hosted the highest magnitude residue dampening across the entire protein.

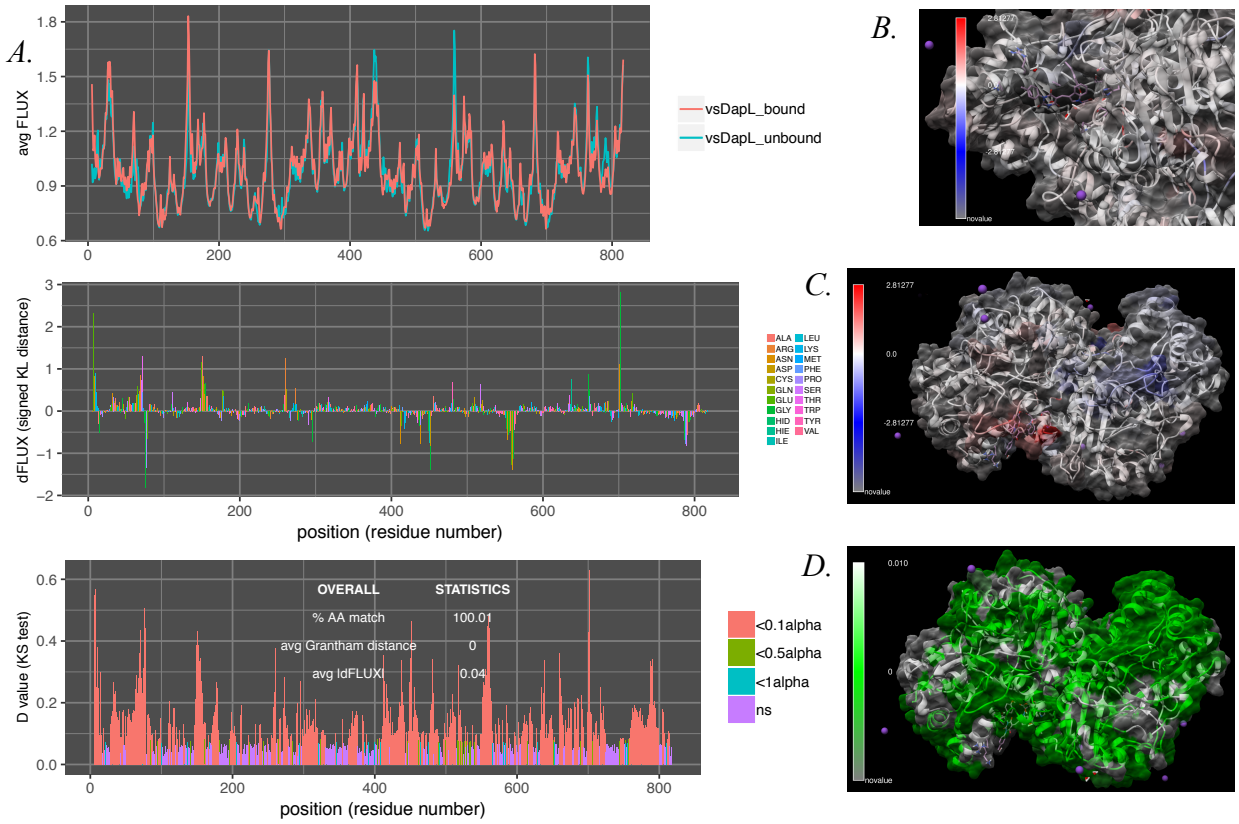
*Figure 31: DROIDS Plot and Images, VsDapL and Rhodanine*




*300 x 0.2ns explicit solvent molecular dynamic simulation of wild type V. spinosum DapL enzyme compared to DapL associated with a rhodanine molecule. (A) Average atom fluctuation profiles (top plot), signed symmetric Kullback-Leibler (KL) divergences in local atom fluctuation distributions, collected from the N, CA, C and O atoms from each amino acid on the polypeptide backbone (middle plot), and P values from a Benjamini-Hochberg corrected Kolmogorov-Smirnov (KS) test indicating significant differences in dynamics is also shown (bottom plot). The KL divergences in local atom fluctuation distributions are shown color mapped to temperature on the protein structure with red indicating softer regions with amplified fluctuation and blue indicating stabilized regions with dampened fluctuation (B) across the entire structure and (C) on the active site. (C) P values from the KS test indicating significant differences in dynamics.*

Statistically significant differences in fluctuations between the unbound VsDapL molecule and VsDapL bound to barbiturate were found upon completion of the DROIDS pipeline (*Figure 32*). The fluctuation (dFLUX) and statistical profiles were asymmetric across the protein, they were not replicated between the two subunits (e.g. the dFLUX of residues 1-10 on chain A was different than the dFLUX of residues 1-10 on chain B). A full list of residues that experienced statistically significant differences in fluctuation are presented in *Figure 38.* Though there was destabilization in all three active site loops in both pockets, there was more reported evidence of statistically significant amplification than dampening. In addition, neither conserved catalytic residue in either pocket was affected.

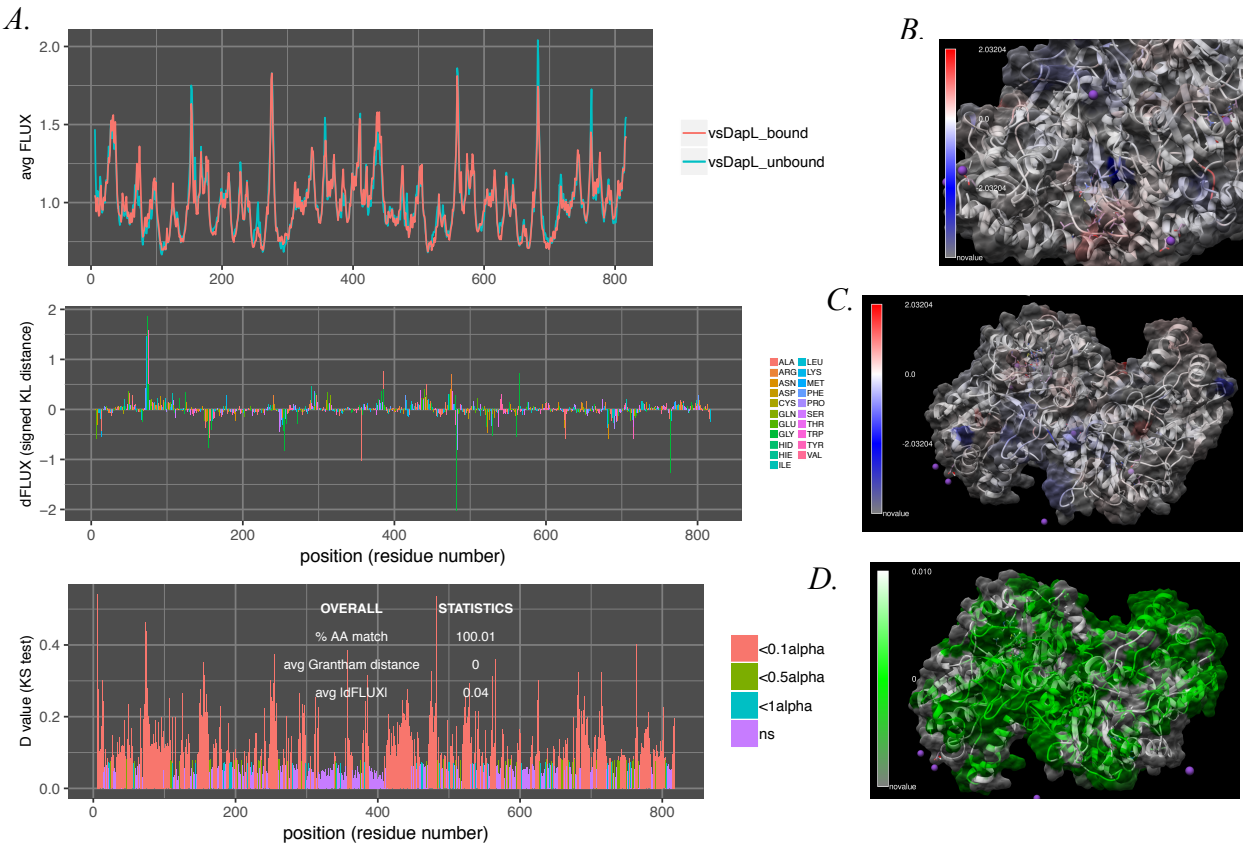
Figure 32: DROIDS Plot and Images, VsDapL and Barbiturate



*300 x 0.2ns explicit solvent molecular dynamic simulation of wild type V. spinosum DapL enzyme compared to DapL associated with a barbiturate molecule. (A) Average atom fluctuation profiles (top plot), signed symmetric Kullback-Leibler (KL) divergences in local atom fluctuation distributions, collected from the N, CA, C and O atoms from each amino acid on the polypeptide backbone (middle plot), and P values from a Benjamini-Hochberg corrected Kolmogorov-Smirnov (KS) test indicating significant differences in dynamics is also shown (bottom plot). The KL divergences in local atom fluctuation distributions are shown color mapped to temperature on the protein structure with red indicating softer regions with amplified fluctuation and blue indicating stabilized regions with dampened fluctuation (B) across the entire structure and (C) on the active site. (C) P values from the KS test indicating significant differences in dynamics.*

Statistically significant differences in fluctuations between the unbound VsDapL molecule and vsDapL bound to thiobarbiturate were found upon completion of the DROIDS pipeline (*Figure 33*). The fluctuation (dFLUX) and statistical profiles were asymmetric across the protein, they were not replicated between the two subunits (e.g. the dFLUX of residues 1-10 on chain A was different than the dFLUX of residues 1-10 on chain B). A full list of residues that experienced statistically significant differences in fluctuation are presented in *Figure 38.* All three loops in the pocket corresponding to the bound molecule were dampened, but only two experienced correlating amplification. In addition, three loops were amplified in the second pocket, but only one experienced correlating dampening. Only one of the two conserved catalytic residues were significantly dampened, the other was unaffected.

*Figure 33: DROIDS Plot and Images, VsDapL and Thiobarbiturate*




*300 x 0.2ns explicit solvent molecular dynamic simulation of wild type V. spinosum DapL enzyme compared to DapL associated with a thiobarbiturate molecule. (A) Average atom fluctuation profiles (top plot), signed symmetric Kullback-Leibler (KL) divergences in local atom fluctuation distributions, collected from the N, CA, C and O atoms from each amino acid on the polypeptide backbone (middle plot), and P values from a Benjamini-Hochberg corrected Kolmogorov-Smirnov (KS) test indicating significant differences in dynamics is also shown (bottom plot). The KL divergences in local atom fluctuation distributions are shown color mapped to temperature on the protein structure with red indicating softer regions with amplified fluctuation and blue indicating stabilized regions with dampened fluctuation (B) across the entire structure and (C) on the active site. (C) P values from the KS test indicating significant differences in dynamics.*

### 8.3.5 DROIDS Summary

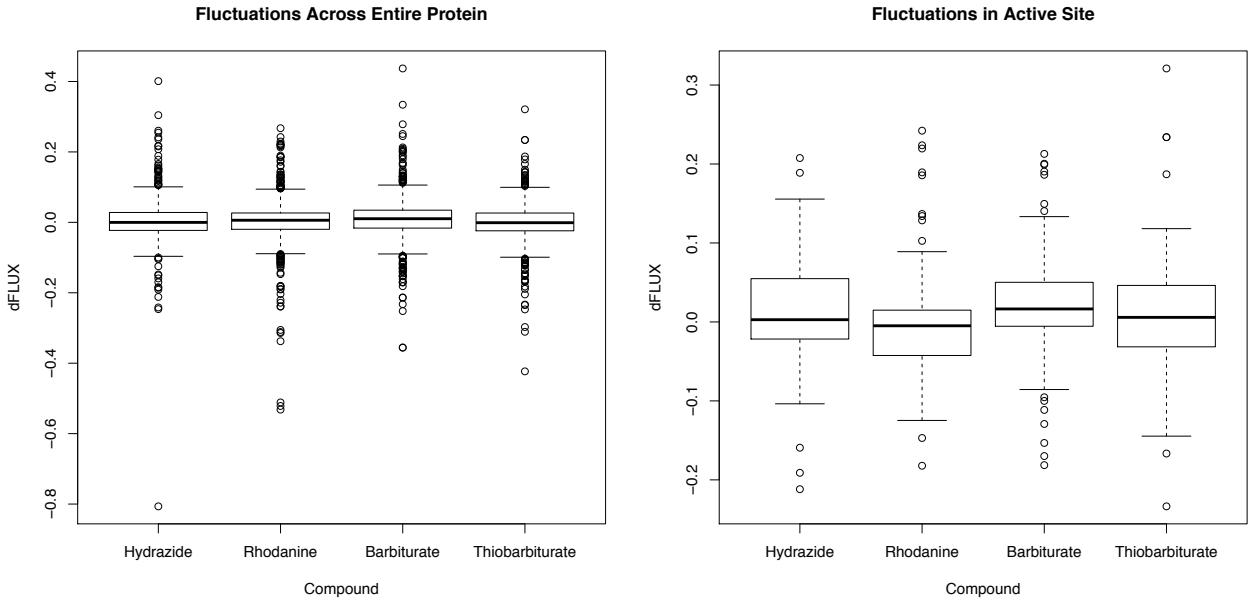
The differences in atomic fluctuation across the bound structure when compared to the unbound structure of each associated compound exhibited comparable distributions of structure-wide significance. The significant differences in fluctuation were clustered around the active sites and dimer interface for each bound structure. The significance profiles and magnitudes of differences in fluctuation across the structures were relatively similar. The distribution of differences in fluctuation across the dapL structures were also relatively similar - dFLUX values ranged from -0.3 to 0.4 in a fairly consistent manner when considering only the values (*Figure 34A*).

The fluctuational differences when localized to the active site, however, indicate the active site of structure bound with rhodanine exhibited more dampening than amplification. The active site of structure bound with hydrazide, though it maintained a fairly wide distribution comparable to that of thiobarbiturate, exhibited a lower mean of difference in fluctuation than thiobarbiturate. The structure bound with barbiturate exhibited the highest magnitude in the distribution of dFLUX values and the dFLUX distribution of the structure bound with thiobarbiturate was the broadest, covering the widest range (*Figure 34B*).

In addition, the number of affected residues across each structure differed with each compound. The structure bound to barbiturate resulted in the most residues exhibiting significant differences in fluctuation (462 residues were affected), followed by thiobarbiturate (417), hydrazide (367), and rhodanine (358) (*Figure 35*).

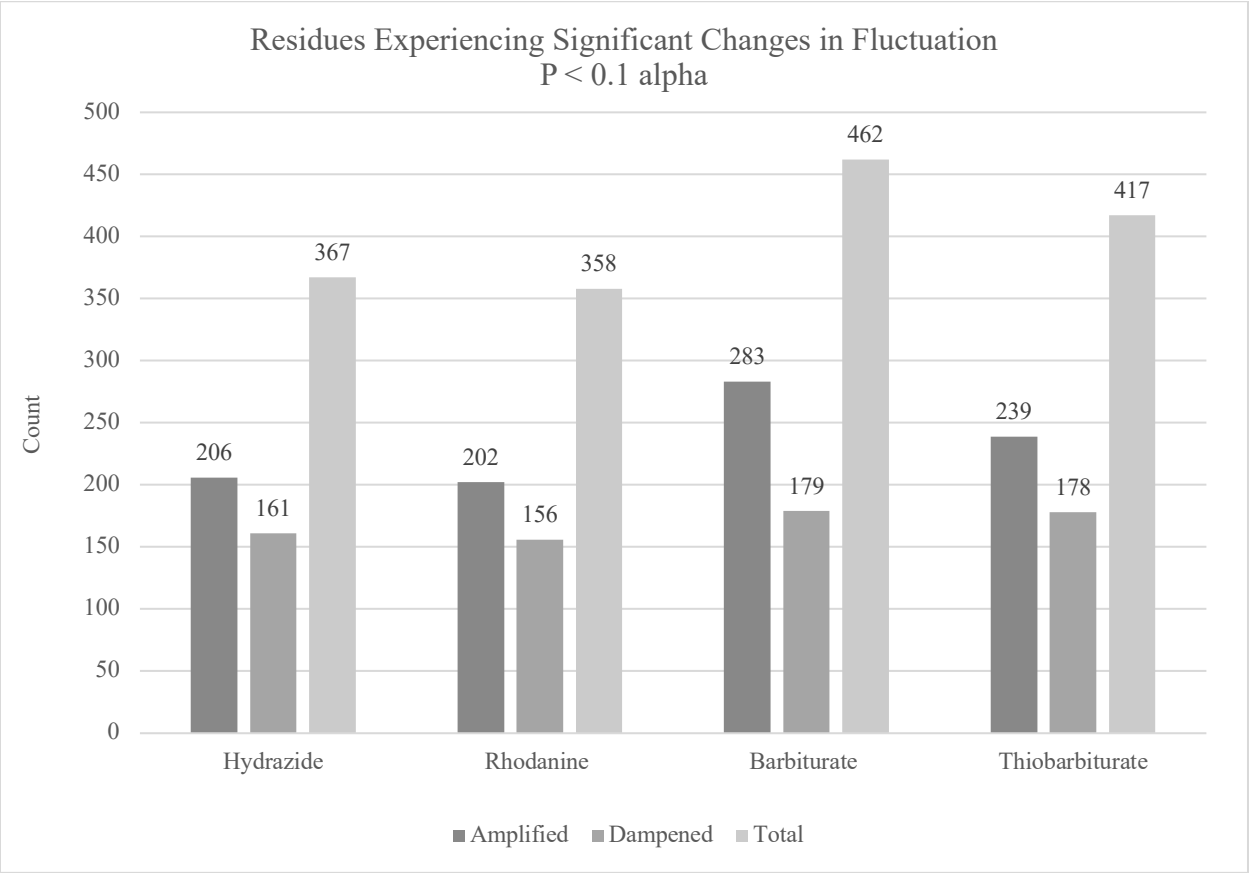
*Figure 34: Atomic Fluctuation Summary*

*A.* *B.*




*Plots constructed in R of (A) the differences in atomic fluctuations across the DapL structure and (B) the differences in atomic fluctuations of active-site residues upon association with each compound when compared to the native unbound structure.*

*Figure 35: Significant fluctuation counts*

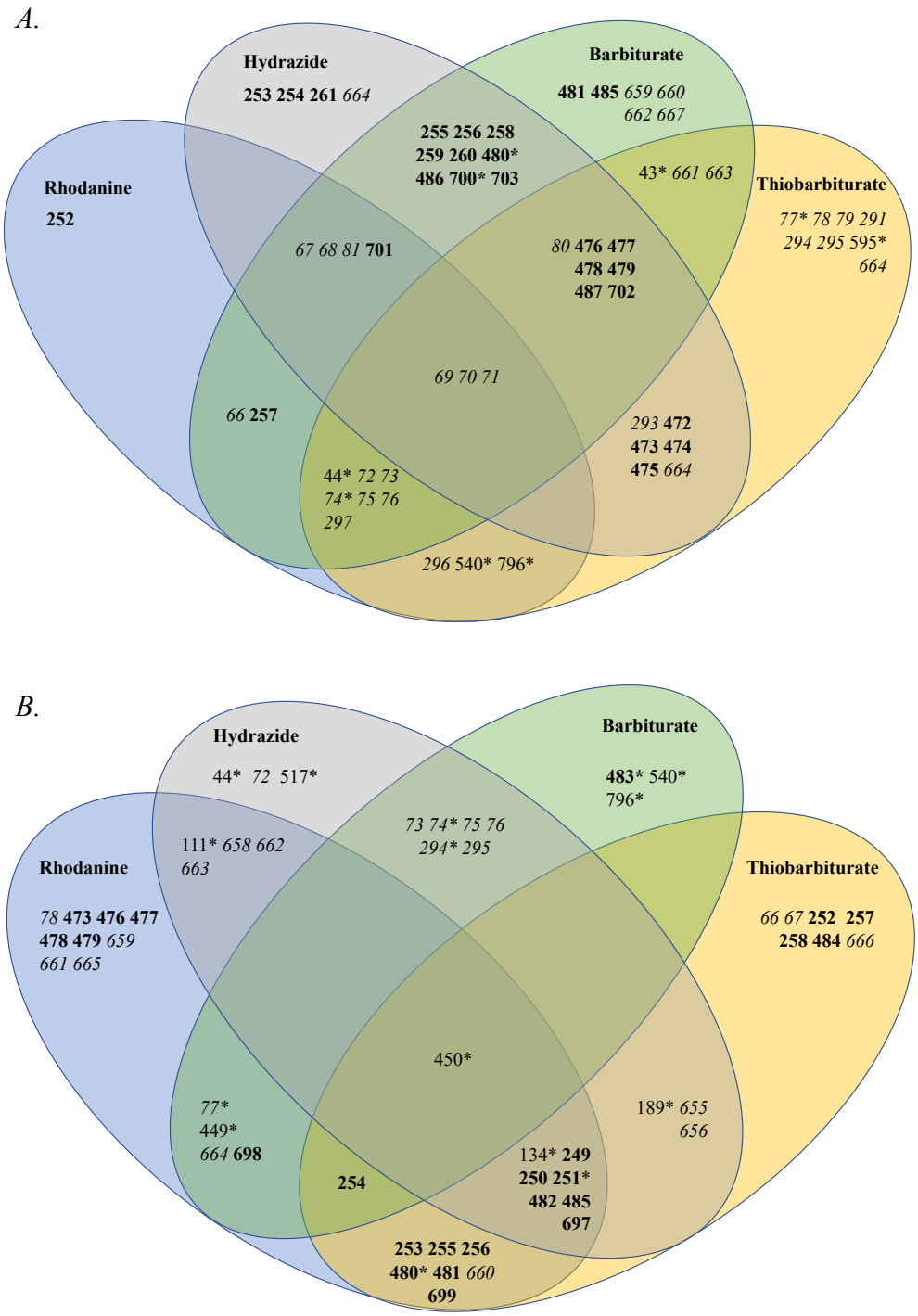



*Numbers of residues in the bound structure experiencing statistically significant differences in fluctuation (less than 0.1 alpha) when compared to the unbound structure for each antagonistic lead compound. Counts of residues experiencing significant amplification, dampening, and the combined total are shown.*

Of the active site residues exhibiting significantly different magnitudes of fluctuation between the bound and unbound structure, only a handful were either consistently amplified or dampened in all four analyses. Residue numbers 69, 70, and 71 were amplified when VsDapL bound to all four inhibitors and residue number 450 was dampened when VsDapL bound to all four inhibitors. There was a higher density of amplified residues when VsDapL bound to hydrazide, barbiturate, and thiobarbiturate and a higher density of dampened residues when VsDapL bound rhodanine and hydrazide. In the structures bound to hydrazide and barbiturate, a large majority of the amplified residues belonged to the active site proximal to the compound docking location, whereas that same pocket was largely dampened in the structures bound to rhodanine and thiobarbiturate (*Figure 36*). As previously mentioned, fluctuations associated with the catalytic lysine of one active site were dampened in the structures bound to hydrazide, rhodanine, and thiobarbiturate. It was unaffected in the structure bound to barbiturate. The fluctuations of atoms associated with long stretches of conserved active site loops were amplified in the structure bound to hydrazide, dampened in the structure bound to rhodanine, and amplified in the structures bound to both barbiturate and thiobarbiturate. Relatively long stretches of conserved active site loops exhibited fluctuation dampening in the structure bound to both hydrazide and thiobarbiturate while short stretches of conserved active site loops exhibited fluctuation amplification in the structure bound to rhodanine and dampening in the structure bound to barbiturate (*Table 5*).

Of the four compounds selected for this study, barbiturate was the best inhibitor of vsDapL *in vitro*, followed by thiobarbiturate, hydrazide, and rhodanine. DapL inhibition by barbiturate and thiobarbiturate was ~10X higher than that by hydrazide, which exhibited ~50X higher inhibition than rhodanine (*Table 7*).

*Figure 36: Summary of Affected Active Site Residues*




*Venn diagrams of residues that were significantly (A) amplified and (B) dampened when comparing the bound and unbound structure for each respective inhibitory compound. Legend:* ***active site 1,*** *active site 2, *Active site residue.*

*Table 5: Summary of Active Site Changes in Fluctuation*

| Compound | Amplified | Dampened |
|---|---|---|
| Hydrazide | *67 68 69 70 71 80 81* **253 254 255 256 258 259 260 261** *293* **472 473 474 475 476 477 478 479 480* 486 487** *664* **700* 701 702 703** | 44* *72 73 74* 75 76* 111* 134* 189* **249 250 251*** *294* 295* 450* **482** 517* *655 656 658 662 663* **697** |
| Rhodanine | 44* *66 67 68 69 70 71 72 73 74* 75 76 81* **252 257** *296 297* 540* **701** 796* | *77* 78* 111* 134* **249 250 251* 253 254 255 256** 449* 450* **473 476 477 478 479 480* 481 482 485** *658 659 660 661 662 663 664 665* **697 698 699** |
| Barbiturate | 43* 44* *66 67 68 69 70 71 80 81* **255 256 257 258 259 260** *297* **476 477 478 479 480* 481 485 486 487** *659 660 661 662 663 667* **700* 701 702 703** | *73 74* 75 76 77** **254** *294* 295* 449* 450* **483*** 540* *664* **698** 796* |
| Thiobarbiturate | 43* 44* *69 70 71 72 73 74* 75 76 77 78 79 80 291 293 294* 295 296 297* **472 473 474 475 476 477 478 479 487** 540* 595* *661 663 664* **702** 796* | *66 67* 134* 189* **249 250 251* 252 253 254 255 256 257 258** 450* **480* 481 482 484 485** *655 656 660 666* **697 699** |

*Lists of residues that exhibited changes in atomic fluctuations, whether amplification or dampening, for each bound structure. Legend:* ***active site 1,*** *active site 2, *Active site residue.*

## 8.4 DISCUSSION

Based on RMSD (root mean square deviation) calculations of the fluctuation of the alpha-carbons of each amino acid from their mean positions, the vsDapL protein is much more stable as a homodimer. The higher and consistently rising RMSD values resulting from the monomer dynamics indicate more instability in the system, while the constant lower RMSD values from the dimer indicate a much more stable structure. It settles into a lowest-energy conformation and maintains that general pattern for the duration of the simulation, as opposed to a constantly fluctuating set of conformations. Based on the sample size and volume of data collected, we have high confidence this trend reflects real-world scenario. In addition, considering the X-ray crystal structures of multiple DapL enzymes and the molecular dynamics results described above, we can conclude that DapL naturally associates into a dimer *in vivo*. As such, all computational biology involving DapL (docking, molecular dynamics, etc.) should include calculations on the homodimer. It is insufficient to assume dynamics for the system based solely on calculations from one subunit. Any conclusions resulting from studies restricted to monomer dynamics are not biologically relevant and cannot be taken in that context. However, because DROIDS is equipped to not only perform the calculations on large, multi-subunit proteins, but also map the statistics at an amino acid resolution, it provides the perfect opportunity for DapL studies.

All antagonistic compounds were successfully docked onto one active site of DapL with a potential energy lower than DapL in its native state and with all potential energies at the same order of magnitude (with rhodanine docking the best, followed by hydrazide and barbiturate, with thiobarbiturate as the least effective). This indicates that all antagonistic compounds would successfully associate with and have the potential to inhibit DapL *in vivo*. However, because DapL is a homodimer, it can be speculated that the inhibitory lead compounds would bind in two

sites. In fact, a test docking of two rhodanine molecules onto DapL indicates better binding affinity for both together than the single molecule. In addition, it docks asymmetrically, perhaps due to allosteric change in the structure (*Figure 38*). For the purpose of DROIDS, compounds were only docked into one active site to accommodate the limitations of pre- and post-calculation processing. Though DROIDS can be applied to multi-subunit proteins, the force field calculations for small molecules, AnteChamber, can only be applied to one molecule per file. Because proper simulations require three files including the file containing the ligand alone in its proper docked conformation, it can be speculated that the molecular dynamics calculations and cpptraj processing would not account for the second ligand. Looking forward, functionality to accommodate multiple docked ligands would be beneficial in the DROIDS software.

As expected, some areas of each protein upon binding were dampened while others experienced an increase in fluctuation. Thus, the effect of an antagonistic lead compound, though identified to inhibit the enzyme *in vitro*, is not always a dampening of protein dynamics. In some cases, the dampening of dynamics could be linked to inhibition, but the commonly seen increase in dynamics could indicate induced instability in the protein upon docking. Both scenarios, theoretically, would alter binding dynamics to and reactions with the substrate to result in the observed inhibition.

Interestingly, the docking results also did not directly reflect either the *in vitro* observed inhibition or the *in silico* molecular dynamics analyses. Barbiturate and hydrazide docked with the same binding affinity, but barbiturate had most significant inhibition while hydrazide was the second least significant. Rhodanine was the worst inhibitor but docked with the best binding affinity. Thiobarbiturate docked with the least binding affinity but was the second most significant inhibitor. The docking position of the molecule and its interaction with neighboring

amino acids also does not necessarily indicate areas of altered molecular dynamics. In most cases, there was significant change in protein dynamics at or proximal to compound position, however some changes in dynamics could not be explained by the local interactions (the changes in protein dynamics near the compound docking location) as they occurred in distant areas of the structure.

The molecular dynamics results support a correlation between inhibitory capacity *in vitro* and overall effect on the protein *in silico*. Binding of the compound that inhibited DapL the most successfully *in vitro* (barbiturate) resulted in the most residues that exhibited significant changes in fluctuation, protein-wide. Binding of the second-most effective compound (thiobarbiturate) resulted in the second-most residues exhibiting significant changes, followed by the third-most effective (hydrazide) and the least effective (rhodanine). This correlation between molecular dynamics and kinetics indicates DROIDS (and more broadly molecular dynamics simulations) can successfully be used in conjunction with laboratory experiments to present both more refined and more broadly impactful results. In addition, DROIDS could theoretically be used prior to any experiments (via high-throughput docking of antagonistic compounds followed by molecular dynamics of the top hits) to save time and cost of laboratory experiments or used to locate local protein regions, domains, or conserved motifs for future study.

The field of molecular dynamics, however, must still be examined with caution as simulations require precise and refined calculations that take into account extremely complex physical characteristics. For example, molecules must be in their lowest potential energy wells for any consistency or meaningful calculations. This task becomes exponentially more difficult as the size of the molecule increases. To date, DapL is the largest protein to be analyzed with DROIDS, and it certainly pushed the limits of the software. If the atomic coordinates were calculated for

longer duration, or sampled for a different number of replicates, it is possible that the refined, amino-acid resolution results would look different. For example, significance results of a 0.5ns simulation differed slightly from those of a 0.2ns simulation for some compounds. In addition, closely-examined active site dynamics for this study do not directly correlate with the *in vitro* study. A rational prediction would indicate the highest active site dampening in the most successful inhibitors, but the opposite was observed *in silico*. Complexes could be ranked as those that maintained larger proportions of and longer stretches of statistically significant residue dampening in both conserved active site residues and active site loops, along with less activation in active site residues and loops, were more effective; the ranking determined by active site dampening did not accurately reflect the *in vitro* data. Examining statistically significant changes in active site molecular dynamics, the structure bound to rhodanine exhibited the highest ratio of active site dampening to amplification, followed by hydrazide and thiobarbiturate, with barbiturate exhibiting the lowest ratio of active site dampening to amplification. This directly contradicts the conclusions drawn from this study and highlights the necessity of analysis repetition, expansion, and validation.

Future work should repeat the analyses detailed above to gather more support for the conclusions and investigate the inconsistencies regarding active site destabilization. The outlying data could be attributed to random sampling, or it could indicate a trend toward a flaw in the analysis. In addition, future work should include investigating the impact of mutations on binding dynamics for insight into the potential for acquired resistance. Mutations that minimize the effect on wild type dynamics but maximize the effect on the binding of the inhibitor would theoretically be the most likely mutations that would contribute to resistance. Overall, though molecular dynamics is still a newly developing field, the potential scope (as indicated by this study) is great.

# 9 CONCLUSIONS

In this study, a combination of wet-bench work and bioinformatics tools was used to gain a comprehensive perspective on the DapL enzyme for evaluation of its potential as a narrow spectrum antibiotic target. Though it is narrowly distributed in bacteria (identified in 13% of all genomes), it is present in a number of clinically important and pathogenic organisms, most notable *Chlamydia trachomatis.* Because of the presence of the DapL pathway in mostly anaerobic, obligate intracellular, or simply difficult to culture organisms, *Verrucomicrobium spinosum* was identified as a putative model for evaluating the essentiality of the DapL enzyme.

However, developing a model system in an environmentally isolated organism, without much prior research or knowledge, proved difficult. Its growth rate and properties in culture were successfully characterized, and standards for genetic manipulation were developed. Though a genetic mutant of the putatively essential *dapL* in *V. spinosum* was not successfully isolated, major strides in assessing the essentiality of the gene and isolating a mutant were made. The differential death rates seen in culture of mutants in supplemented liquid media, the inability to culture anything on supplemented solid media, and the microscopic images taken of the different putative mutants, when combined, support the essentiality of *dapL*. In fact, the gene may be so essential and its loss so detrimental to the organism that the recovery of a mutant may be impossible. The loss of both lysine in protein synthesis and *meso*-DAP in peptidoglycan crosslinking may cause deficiencies and failures significant enough to cause instant cell death upon gene interruption or replacement. Alternatively, the inability to isolate a mutant could also be attributed to a flaw in methodology – whether it is unsuccessful transformations or improper recovery and isolation.

To supplement the wet-bench analyses, comparative molecular dynamics simulations of the *V.* spinosum DapL enzyme and four antagonistic lead compounds were performed. The success in correlating whole protein destabilization identified *in silico* with inhibition identified *in vitro* supports the efficacy of using such a tool in conjunction with or prior to laboratory experiments to refine compound identification and analysis. However, the volume and ambiguity in the data highlights the need for verification and validation of any large-scale analyses via wet-bench experiments. Proper care in maintaining the biologically relevant perspective must be taken to ensure the data is not skewed, misinterpreted, or simply wrong. When proper steps are taken, powerful and elegant conclusions can be drawn about the dynamic behavior, essentiality, and interactions of proteins and enzymes of interest. In the case of this study, though the essentiality of DapL was not explicitly proven, support for its essentiality was found and the effect of four antagonistic lead compounds on the dynamic behavior of the enzyme was successfully characterized.

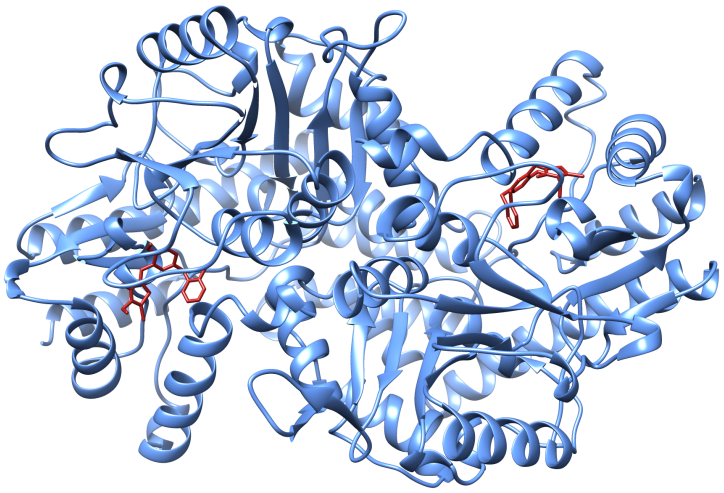
# 10 SUPPLEMENTAL FIGURES

*Table 6: RGI CARD Scan*

| | |
|---|---|
| Contig | ABIZ01000001.1_4980 |
| Start | 6205457 |
| Stop | 6206245 |
| Homology | Loose |
| Term | catB8 |
| % ID | 32.69 |
| Drug Class | phenicol antibiotic |
| Resistance Mechanism | antibiotic inactivation |
| AMR Family | chloramphenicol acetyltransferase (CAT) |

*The identification of a gene in V. spinosum homologous to a gene conferring chloramphenicol resistance.*

*Figure 37: Rhodanine Double Docking*



| Ligand | Binding Affinity (kcal/mole) | rmsd/ub | rmsd/lb |
|---|---|---|---|
| VsDapL_1A | -8.5 | 0 | 0 |
| VsDapL_2A | -8.1 | 43.936 | 40.674 |
| VsDapL_3A | -8.1 | 3.163 | 2.227 |
| VsDapL_4A | -8 | 2.248 | 1.788 |
| VsDapL_5A | -7.9 | 18.639 | 16.279 |
| VsDapL_6A | -7.8 | 46.198 | 43.602 |
| VsDapL_1B | -8.4 | 0 | 0 |
| VsDapL_2B | -7.5 | 46.749 | 44.442 |
| VsDapL_3B | -7.2 | 40.06 | 38.123 |
| VsDapL_4B | -7.2 | 33.441 | 32.209 |
| VsDapL_5B | -7.1 | 33.272 | 31.882 |
| VsDapL_6B | -7 | 37.817 | 36.146 |

*(A) The structure of two rhodanine molecules (red) associated with the VsDapL enzyme (blue). (B) Docking results for the double rhodanine and VsDapL association. The lowest energy structure is reported first for each molecule (A and B), and the RMSD of the following structures is reported as compared to the position of the first conformation.*

*Figure 38: Full list of dampened and activated residues for each MD run*

*A. Hydrazide*

| Amplified | Dampened |
|---|---|
| 5 6 36 37 41 42 45 46 47 48 49 50 51 52 53 54 55 56 57 58 67 68 69 70 71 80 81 82 85 86 89 90 92 93 94 95 96 102 103 104 105 106 107 108 128 198 207 208 209 227 228 229 230 231 232 233 234 235 236 237 238 239 240 247 253 254 255 256 258 259 260 261 262 263 267 284 285 286 287 293 311 312 313 314 315 316 317 318 319 320 321 322 323 325 326 355 356 357 393 394 395 396 397 398 399 400 401 402 403 405 411 414 415 417 418 419 420 424 446 447 448 459 460 463 467 468 471 472 473 474 475 476 477 478 479 480 486 487 489 495 496 497 503 504 505 512 513 514 530 535 536 542 543 544 546 547 550 551 552 555 556 557 558 559 560 561 562 567 568 577 578 582 583 584 603 604 605 606 611 616 617 624 664 677 678 679 680 681 682 683 685 686 700* 701 702 703 705 707 721 742 761 762 763 800 801 802 | 11 12 14 15 16 19 22 23 26 27 30 31 32 33 44 63 72 73 74 75 76 98 110 111 112 113 114 115 116 117 118 119 120 121 122 133 134 135 136 138 139 140 142 143 149 150 151 152 153 155 156 157 158 167 168 169 172 173 175 176 178 179 185 186 187 188 189 210 217 218 219 220 221 222 223 224 249 250 251 269 270 271 272 273 274 275 276 277 278 294 295 348 358 375 381 382 383 384 409 410 413 422 442 443 450 451 452 453 454 455 482 516 517 518 519 520 521 522 523 524 525 526 527 528 529 532 538 554 564 570 576 590 591 598 599 614 637 653 654 655 656 658 662 663 668 669 691 693 694 697 709 716 785 786 788 789 790 791 792 815 816 |

*B. Rhodanine*

| Amplified | Dampened |
|---|---|
| 44 45 47 64 65 66 67 68 69 70 71 72 73 74 75 76 81 82 83 85 86 89 90 91 98 99 101 102 106 108 115 163 166 167 168 169 170 171 172 173 186 187 188 192 193 194 226 240 241 252 257 265 266 267 268 269 271 272 273 274 275 277 278 282 285 286 287 288 289 296 297 308 309 310 311 312 314 315 339 340 341 342 351 352 353 354 386 387 411 413 414 415 416 417 418 419 420 421 422 423 432 433 434 435 436 437 438 439 440 441 442 444 445 446 453 471 494 495 496 497 498 499 500 501 502 503 504 505 511 512 519 530 536 538 539 540 541 542 544 546 547 548 549 550 551 552 553 569 577 578 579 582 583 592 593 600 601 602 625 626 627 628 629 630 631 641 677 678 679 | 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 41 53 54 57 58 77* 78 109 110 111 112 113 116 117 118 119 122 123 124 130 133 134 146 150 151 152 153 154 155 176 177 178 210 212 247 249 250 251 253 254 255 256 299 300 302 303 304 356 357 381 394 398 409 448 449 450 451 452 458 459 460 462 463 464 467 473 476 477 478 479 480 481 482 485 514 515 520 521 522 532 555 556 557 558 559 560 561 562 563 564 609 633 636 637 658 659 660 661 662 663 664 665 687 688 689 690 691 692 695 696 697 698 699 742 743 744 745 761 762 763 764 799 801 |

| | |
|---|---|
| 680 681 682 683 701 710 714 715 716 717 718<br>719 720 721 722 723 724 730 731 733 734 750<br>751 752 756 771 774 775 781 783 795 796 816 | |

*C. Barbiturate*

| Amplified | Dampened |
|---|---|
| 5 6 7 8 9 10 11 12 27 28 29 30 31 32<br>33 34 35 36 37 38 39 40 41 43 44 45 50<br>51 52 53 54 55 56 57 58 59 60 61 62 63 64<br>65 66 67 68 69 70 71 80 81 82 83 85 86<br>91 109 110 114 124 130 132 145 146 147 148 149<br>150 151 152 153 154 155 156 157 158 159 161<br>165 169 176 177 178 179 210 211 212 221 222<br>231 248 255 256 257 258 259 260 267 270 271<br>272 273 274 277 297 298 299 300 301 302 303<br>304 305 306 307 308 309 310 311 312 313 314<br>315 316 317 318 319 324 325 326 327 328 329<br>344 353 354 362 363 364 365 366 367 368 369<br>370 371 372 373 374 382 383 406 409 410 414<br>454 455 456 457 458 459 460 462 465 466 467<br>470 471 476 477 478 479 480 481 485 486 487<br>488 496 504 505 506 507 508 509 510 511 512<br>513 516 518 519 536 568 569 570 573 574 578<br>584 585 586 588 589 601 602 605 606 607 608<br>609 610 611 612 613 616 617 618 619 620 621<br>632 633 634 635 636 637 638 639 640 641 642<br>643 644 645 646 647 648 649 651 659 660 661<br>662 663 667 668 669 670 671 672 673 674 675<br>676 680 681 682 683 684 685 700 701 702 703<br>704 707 708 709 710 711 712 716 717 718 719<br>720 721 722 723 803 804 805 806 | 13 14 15 17 18 48 73 74 75 76 77 95 96<br>97 98 112 115 116 117 118 120 122 171 172 173<br>174 175 194 196 197 198 199 201 202 203 237<br>238 239 240 254 262 265 266 275 276 281 282<br>289 290 294 295 387 411 412 413 416 417 418<br>419 421 422 423 424 425 426 427 428 429 430<br>431 432 433 434 435 436 437 438 439 440 441<br>443 444 446 448 449 450 451 452 453 483 497<br>498 500 501 502 521 526 529 540 541 542 547<br>548 549 550 551 552 553 554 555 556 557 558<br>559 560 561 562 563 564 596 597 627 664 689<br>690 691 692 698 714 745 746 747 748 749 757<br>758 759 760 761 762 763 764 765 766 767 768<br>769 770 771 772 773 774 775 776 777 778 779<br>780 781 782 783 784 785 786 787 788 789 790<br>791 792 793 794 795 796 797 798 811 812 813 |

*D. Thiobarbiturate*

| Amplified | Dampened |
|---|---|
| 16 19 32 33 34 35 36 37 39 40 42 43 44<br>47 48 49 50 51 53 54 69 70 71 72 73 74<br>75 76 77 78 79 80 81 82 83 84 85 86 87 88<br>89 90 91 92 93 94 95 96 97 99 100 101 102<br>103 104 105 106 107 115 123 124 125 146 176<br>177 178 234 235 239 240 241 266 267 268 274 | 5 6 7 8 9 10 11 12 13 25 26 27 41 62<br>63 65 66 67 110 122 126 128 129 130 132 133<br>134 147 148 149 150 151 152 153 154 155 156<br>157 158 159 160 161 168 169 170 171 172 179<br>186 187 188 189 190 196 199 200 201 212 215<br>217 218 219 225 226 227 228 229 246 247 248 |

| | |
|---|---|
| 287 288 289 290 291 293 294 295 296 297 305<br>310 311 312 314 315 338 339 353 358 359 360<br>361 362 363 378 379 381 382 383 384 385 386<br>409 413 414 415 420 421 422 423 424 425 426<br>427 428 429 430 431 432 433 434 435 436 437<br>438 439 440 441 442 443 444 445 446 448 452<br>470 471 472 473 474 475 476 477 478 479 487<br>488 501 502 503 504 505 509 513 518 539 540<br>541 542 551 552 563 564 565 573 574 575 576<br>581 582 595 596 597 598 599 600 614 615 616<br>637 646 647 653 661 663 664 674 675 678 702<br>704 705 725 726 737 740 741 742 743 744 745<br>746 747 753 758 769 770 771 772 773 774 775<br>776 777 778 779 780 781 782 783 786 791 792<br>794 795 796 797 798 799 800 801 802 803 | 249 250 251 252 253 254 255 256 257 258 270<br>271 276 277 278 356 357 411 416 417 418 419<br>450 451 455 456 457 480 481 482 484 485 493<br>494 497 519 520 521 522 523 524 525 526 527<br>528 529 531 535 536 545 546 558 559 560 561<br>562 587 621 622 623 624 625 626 654 655 656<br>660 666 680 681 682 685 686 687 688 689 690<br>691 692 693 694 695 696 697 699 708 709 710<br>711 712 713 714 715 716 717 718 719 720 761<br>762 763 764 787 788 789 814 815 816 |

*Numbers of vsDapL residues that experienced significant changes in atomic fluctuation when bound with (A) hydrazide (B) rhodanine (C) barbiturate and (D) thiobarbiturate.*

*Table 7: VsDapL Inhibition*

| Compound | $IC_{50}$ Value |
|---|---|
| Hydrazide | 47 |
| Rhodanine | 250 |
| Barbiturate | 4.7 |
| Thiobarbiturate | 5.7 |

*Inhibition of the VsDapL enzyme as defined by $IC_{50}$ value for the four compounds used in this study. Adapted from (McKinnie et al., 2014).*

# 11 REFERENCES


Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). General Recombination. *Molecular Biology of the Cell. 4th Edition*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK26898/

Ansdell, V. (2012). 70 - Leptospirosis. In E. C. Jong & D. L. Stevens (Eds.), *Netter's Infectious Diseases* (pp. 425–429). https://doi.org/10.1016/B978-1-4377-0126-5.00070-7

Babbitt, G. A., Mortensen, J. S., Coppola, E. E., Adams, L. E., & Liao, J. K. (2018). DROIDS 1.20: A GUI-Based Pipeline for GPU-Accelerated Comparative Protein Dynamics. *Biophysical Journal*, *114*(5), 1009–1017. https://doi.org/10.1016/j.bpj.2018.01.020

Berry, C. D., Hooton, T. M., Collier, A. C., & Lukehart, S. A. (1987). Neurologic Relapse after Benzathine Penicillin Therapy for Secondary Syphilis in a Patient with HIV Infection. *New England Journal of Medicine*, *316*(25), 1587–1589. https://doi.org/10.1056/NEJM198706183162507

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., … Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, *42*(W1), W252–W258. https://doi.org/10.1093/nar/gku340

Burstain, J. M., Grimprel, E., Lukehart, S. A., Norgard, M. V., & Radolf, J. D. (1991). Sensitive detection of Treponema pallidum by using the polymerase chain reaction. *Journal of Clinical Microbiology*, *29*(1), 62–69.

CASE, D. A., CHEATHAM, T. E., DARDEN, T., GOHLKE, H., LUO, R., MERZ, K. M., … WOODS, R. J. (2005). The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry*, *26*(16), 1668–1688. https://doi.org/10.1002/jcc.20290

ChemBridge | Home. (n.d.). Retrieved December 14, 2018, from https://www.chembridge.com/

Chylinski, K., Le Rhun, A., & Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biology*, *10*(5), 726–737. https://doi.org/10.4161/rna.24321

Dallakyan, S., & Olson, A. J. (2015). Small-Molecule Library Screening by Docking with PyRx. In J. E. Hempel, C. H. Williams, & C. C. Hong (Eds.), *Chemical Biology: Methods and Protocols* (pp. 243–250). https://doi.org/10.1007/978-1-4939-2269-7_19

Dedysh, S. N., Pankratov, T. A., Belova, S. E., Kulichevskaya, I. S., & Liesack, W. (2006). Phylogenetic analysis and in situ identification of bacteria community composition in an acidic Sphagnum peat bog. *Applied and Environmental Microbiology*, *72*(3), 2110–2117. https://doi.org/10.1128/AEM.72.3.2110-2117.2006

Dobson, R. C. J., Girón, I., & Hudson, A. O. (2011). L,L-Diaminopimelate Aminotransferase from Chlamydomonas reinhardtii: A Target for Algaecide Development. *PLOS ONE*, *6*(5), e20439. https://doi.org/10.1371/journal.pone.0020439

Domman, D. B., Steven, B. T., & Ward, N. L. (2011). Random transposon mutagenesis of Verrucomicrobium spinosum DSM 4136T. *Archives of Microbiology*, *193*(4), 307–312. https://doi.org/10.1007/s00203-010-0666-5

Finegold, S., & Sussman, M. (2002). Anaerobic InfectionsA Clinical Overview. In *Molecular Medical Microbiology* (Vol. 3, pp. 1867–1874). https://doi.org/10.1016/B978-012677530-3/50307-X

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., … Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67–71. https://doi.org/10.1038/nature09523

Griffiths, E., & Gupta, R. S. (2007). Phylogeny and shared conserved inserts in proteins provide evidence that Verrucomicrobia are the closest known free-living relatives of chlamydiae. *Microbiology*, *153*(8), 2648–2654. https://doi.org/10.1099/mic.0.2007/009118-0

Hernández-Santoyo, A., Tenorio-Barajas, A. Y., Altuzar, V., Vivanco-Cid, H., & Mendoza-Barrera, C. (2013). Protein-Protein and Protein-Ligand Docking. *Protein Engineering - Technology and Application*. https://doi.org/10.5772/56376

Hudson, A. O., Bless, C., Macedo, P., Chatterjee, S. P., Singh, B. K., Gilvarg, C., & Leustek, T. (2005). Biosynthesis of lysine in plants: evidence for a variant of the known bacterial pathways. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1721*(1), 27–36. https://doi.org/10.1016/j.bbagen.2004.09.008

Hudson, A. O., Singh, B. K., Leustek, T., & Gilvarg, C. (2006). An ll-Diaminopimelate Aminotransferase Defines a Novel Variant of the Lysine Biosynthesis Pathway in Plants. *Plant Physiology*, *140*(1), 292–301. https://doi.org/10.1104/pp.105.072629

Hudson, Klartag, A., Gilvarg, C., Dobson, R. C. J., Marques, F. G., & Leustek, T. (2011). Dual diaminopimelate biosynthesis pathways in Bacteroides fragilis and Clostridium

thermocellum. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1814*(9), 1162–1168. https://doi.org/10.1016/j.bbapap.2011.04.019

Hutton, C. A., Perugini, M. A., & Gerrard, J. A. (2007). Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Molecular BioSystems*, *3*(7), 458–465. https://doi.org/10.1039/B705624A

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–5433.

Jansen, R., Embden, J. D. A. van, Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), 1565–1575.

Jensen, F. (2007). *Introduction to computational chemistry* (2nd ed). Chichester, England ; Hoboken, NJ: John Wiley & Sons.

Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology*, *31*(3), 233–239. https://doi.org/10.1038/nbt.2508

Lewars, E. (2003). Computational Chemistry - Introduction to the Theory and Applications of Molecular and Quantum Mechanics. *COMPUTATIONAL CHEMISTRY*, 482.

Liang, X., Potter, J., Kumar, S., Ravinder, N., & Chesnut, J. D. (2017). Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9

nuclease, and donor DNA. *Journal of Biotechnology*, *241*, 136–146. https://doi.org/10.1016/j.jbiotec.2016.11.011

Liepman, A. H., & Olsen, L. J. (2004). Genomic Analysis of Aminotransferases in Arabidopsis thaliana. *Critical Reviews in Plant Sciences*, *23*(1), 73–89. https://doi.org/10.1080/07352680490273419

Lindsay, K. W., Bone, I., & Fuller, G. (2010). SECTION V - MULTIFOCAL NEUROLOGICAL DISEASE AND ITS MANAGEMENT. In K. W. Lindsay, I. Bone, & G. Fuller (Eds.), *Neurology and Neurosurgery Illustrated (Fifth Edition)* (pp. 489–563). https://doi.org/10.1016/B978-0-443-06957-4.50011-2

Mabey, D. (2008). Trachoma: Recent Developments. In A. Finn & A. J. Pollard (Eds.), *Hot Topics in Infection and Immunity in Children IV* (pp. 98–107). Springer New York.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., … Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, *339*(6121), 823–826. https://doi.org/10.1126/science.1232033

McKinnie, S. M. K., Rodriguez-Lopez, E. M., Vederas, J. C., Crowther, J. M., Suzuki, H., Dobson, R. C. J., … Hudson, A. O. (2014). Differential response of orthologous l,l-diaminopimelate aminotransferases (DapL) to enzyme inhibitory antibiotic lead compounds. *Bioorganic & Medicinal Chemistry*, *22*(1), 523–530. https://doi.org/10.1016/j.bmc.2013.10.055

Mishori, R., McClaskey, E. L., & WinklerPrins, V. (2012). *Chlamydia Trachomatis* Infections: Screening, Diagnosis, and Management. *American Family Physician*, *86*(12), 1127–1132.

Nachar, V. R., Savka, F. C., McGroty, S. E., Donovan, K. A., Dobson, R. C. J., North, R. A., … Hudson, A. O. (2012). Genomic and Biochemical Analysis of the Diaminopimelate and Lysine Biosynthesis Pathway in Verrucomicrobium spinosum: Identification and Partial Characterization of L,L-Diaminopimelate Aminotransferase and UDP-N-Acetylmuramoylalanyl-D-glutamyl-2,6-meso-Diaminopimelate Ligase. *Frontiers in Microbiology*, *3*. https://doi.org/10.3389/fmicb.2012.00183

Nishida, H., Nishiyama, M., Kobashi, N., Kosuge, T., Hoshino, T., & Yamane, H. (1999). A prokaryotic gene cluster involved in synthesis of lysine through the amino adipate pathway: a key to the evolution of amino acid biosynthesis. *Genome Research*, *9*(12), 1175–1183.

Packiam, M., Weinrick, B., Jacobs, W. R., & Maurelli, A. T. (2015). Structural characterization of muropeptides from Chlamydia trachomatis peptidoglycan by mass spectrometry resolves “chlamydial anomaly.” *Proceedings of the National Academy of Sciences of the United States of America*, *112*(37), 11660–11665. https://doi.org/10.1073/pnas.1514026112

Pallen, M. J., Beatson, S. A., & Bailey, C. M. (2005). Bioinformatics, genomics and evolution of non-flagellar type-III secretion systems: a Darwinian perpective. *FEMS Microbiology Reviews*, *29*(2), 201–229. https://doi.org/10.1016/j.femsre.2005.01.001

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pilhofer, M., Aistleitner, K., Biboy, J., Gray, J., Kuru, E., Hall, E., … Jensen, G. J. (2013). Discovery of chlamydial peptidoglycan reveals bacteria with murein sacculi but without FtsZ. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms3856

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, *9*(7), 3084–3095. https://doi.org/10.1021/ct400341p

Sadowski, J., Gasteiger, J., & Klebe, G. (1994). Comparison of Automatic Three-Dimensional Model Builders Using 639 X-ray Structures. *Journal of Chemical Information and Computer Sciences*, *34*(4), 1000–1008. https://doi.org/10.1021/ci00020a039

Sangwan, P., Chen, X., Hugenholtz, P., & Janssen, P. H. (2004). Chthoniobacter flavus gen. nov., sp. nov., the First Pure-Culture Representative of Subdivision Two, Spartobacteria classis nov., of the Phylum Verrucomicrobia. *Applied and Environmental Microbiology*, *70*(10), 5875–5881. https://doi.org/10.1128/AEM.70.10.5875-5881.2004

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Research*, *39*(21), 9275–9282. https://doi.org/10.1093/nar/gkr606

Schwab, C. H. (2010). Conformations and 3D pharmacophore searching. *Drug Discovery Today: Technologies*, *7*(4), e245–e253. https://doi.org/10.1016/j.ddtec.2010.10.003

Triassi, A. J., Wheatley, M. S., Savka, M. A., Gan, H. M., Dobson, R. C. J., & Hudson, A. O. (2014). L,L-diaminopimelate aminotransferase (DapL): a putative target for the development of narrow-spectrum antibacterial compounds. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00509

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. https://doi.org/10.1002/jcc.21334

TypeOne™ Restriction Inhibitor. (n.d.). Retrieved May 30, 2018, from https://www.lucigen.com/TypeOne-Restriction-Inhibitor/

Velasco, A. M., Leguina, J. I., & Lazcano, A. (2002). Molecular Evolution of the Lysine Biosynthetic Pathways. *Journal of Molecular Evolution*, *55*(4), 445–449. https://doi.org/10.1007/s00239-002-2340-2

Velick, S. F., & Vavra, J. (1962). A Kinetic and Equilibrium Analysis of the Glutamic Oxaloacetate Transaminase Mechanism. *Journal of Biological Chemistry*, *237*(7), 2109–2122.

Ventola, C. L. (2015). The Antibiotic Resistance Crisis. *P T*, *40*(4), 277–283.

Vergin, K. L., Urbach, E., Stein, J. L., DeLong, E. F., Lanoil, B. D., & Giovannoni, S. J. (1998). Screening of a fosmid library of marine environmental genomic DNA fragments reveals four clones related to members of the order Planctomycetales. *Applied and Environmental Microbiology*, *64*(8), 3075–3078.

Wagner, M., & Horn, M. (2006). The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. *Current Opinion in Biotechnology*, *17*(3), 241–249. https://doi.org/10.1016/j.copbio.2006.05.005

Welcome to MN-AM | MN-AM. (n.d.). Retrieved December 14, 2018, from https://www.mn-am.com/

Wexler, H. M. (2007). Bacteroides: the Good, the Bad, and the Nitty-Gritty. *Clinical Microbiology Reviews*, *20*(4), 593–621. https://doi.org/10.1128/CMR.00008-07

Wu, X., Ren, G., & Huntley, J. F. (2015). Generating Isogenic Deletions (Knockouts) in Francisella tularensis, a Highly-infectious and Fastidious Gram-negative Bacterium. *Bio-Protocol*, *5*(12), e1500.